

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:27:37 ; Search time 21900 Seconds  
(without alignments)  
11635.893 Million cell updates/sec

Title: US-09-913-159a-10

Perfect score: 6229

Sequence: 1 ctgacgcgcgcctctagcgcgc.....attcccccgaagtcgcac 6229

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_stcs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6229	100.0	6229	6 AX032758	AX032758 Sequence
2	2676.2	43.0	6229	6 AR067813	AR067813 Sequence
3	2676.2	43.0	9709	6 AR224437	AR224437 Sequence
4	2676.2	43.0	9709	6 AX032749	AX032749 Sequence
5	2676.2	43.0	9709	6 HIVNL43	HIVNL43 Sequence
6	2676.2	43.0	14824	12 AF324493	AF324493 HIV-1 vec
7	2665	42.8	9000	14 HIVUD6942	HIVUD6942 Sequence
8	2610.6	41.9	9754	14 HIV2132	HIV2132 Sequence
9	2607.4	41.9	9699	14 AF070521	AF070521 HIV-1 E9
10	2599.4	41.7	9213	6 E00987	E00987 Genomic DNA
11	2599.4	41.7	9213	6 I04549	I04549 Sequence 11
12	2599.4	41.7	9752	14 HIVMCK1	HIVMCK1 Sequence
13	2599.4	41.7	9770	14 HIVPV22	HIVPV22 Sequence
14	2598.2	41.7	3156	6 E01088	E01088 Nucleic ac1
15	2597.8	41.7	8932	14 HIVBHL02	HIVBHL02 Sequence
16	2597.8	41.7	8933	6 AR094659	AR094659 Sequence
17	2597.8	41.7	8933	6 AX078307	AX078307 Sequence
18	2597.8	41.7	8933	6 AX078308	AX078308 Sequence
19	2597.8	41.7	8933	6 AX078313	AX078313 Sequence
20	2597.8	41.7	9748	14 REHTLV3	REHTLV3 Sequence
21	2597.8	41.7	9749	6 I07983	I07983 Sequence 1
22	2596.2	41.7	9748	6 E01099	E01099 DNA sequence
23	2595	41.7	3156	14 HIVHXB3	HIVHXB3 Sequence
24	2593	41.6	9181	6 AX074066	AX074066 Sequence
25	2593	41.6	9181	6 AF033819	AF033819 HIV-1, CO
26	2593	41.6	9719	6 AR268768	AR268768 Sequence
27	2593	41.6	9719	14 HIVHXB2CG	HIVHXB2CG Sequence
28	2583.4	41.5	9795	14 HIVTH475A	HIVTH475A Sequence
29	2574.4	41.3	9193	6 A04321	A04321 Human immun
30	2574.4	41.3	9229	14 HIVBRUCG	HIVBRUCG Sequence
31	2572.4	41.3	9193	6 A07867	A07867 Human immun
32	2572.2	41.3	9609	14 HIVU12055	HIVU12055 Sequence
33	2563.4	41.2	2565	6 BD000747	BD000747 Recombembl
34	2560.4	41.1	2565	6 AX166280	AX166280 Sequence
35	2558.4	41.1	12479	6 AR287304	AR287304 Sequence
36	2558.4	41.1	12494	6 AR140325	AR140325 Sequence
37	2558.4	41.1	12494	6 BD105983	BD105983 Animal mo
38	2554.6	41.0	3563	6 AR094661	AR094661 Sequence
39	2554.6	41.0	3563	14 HIVH3BH8	HIVH3BH8 Sequence
40	2553	41.0	9781	14 HIVF12CG	HIVF12CG Sequence
41	2543.6	40.8	2918	6 I05801	I05801 Sequence 2
42	2543.4	40.8	9091	6 A00647	A00647 Lymphadenop
43	2543.4	40.8	9091	6 BD165833	BD165833 DNA obtai
44	2543.4	40.8	9094	6 A10447	A10447 Complete sy
45	2518.6	40.4	2565	14 HIVGPL60BN	HIVGPL60BN Sequence

#### ALIGNMENTS

RESULT 1	AX032758	6229 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX032758				
DEFINITION	Sequence 10 from Patent WO0047223.				
ACCESSION	AX032758				
VERSION	AX032758.1	GI:10279735			
KEYWORDS					
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1				
AUTHORS	Schreiber, M.				
TITLE	Viral vaccine				
JOURNAL	Patent: WO 0047223-A 10 17-AUG-2000;				
	SCHREIBER MICHAEL (DE) ; STRATHMANN AG & CO (DE)				

FEATURES		Location/Qualifiers
Source		
	1. 6229	/organism="synthetic construct"
		/mol_type="genomic DNA"
		/db_xref="taxon:32630"
		/note="Synthetische"
	sig_peptide	1293. 1295
		/note="envy ANG"
	misc_feature	1377. 1379
		/note="envy AGT, gp120 Anfang"
	misc_feature	1397. 1403
		/note="BseEII-Schmittstelle"
	misc_feature	3537. 3542
		/note="BamHI-Schmittstelle"
	misc_feature	3855. 3857
		/note="envy TAA, Stop"
BASE COUNT	1795 a 1332 c 1518 g 1584 t	
ORIGIN		
Query Match	100.0%; Score 6229; DB 6; Length 6229;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 6229; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 CTGACGGGCGCTGTAGGCGGCGCATTAAGCGCGGGGTGTGATGGTTACGCCACCGTGA	60
Db	1 CTGACGGGCGCTGTAGGCGGCGCATTTAAGCGCGGGGTGTGATGGTTACGCCACCGTGA	60
OY	61 CCGGTACACTTGGCAGCGCCCTTAGCGCCGCTCTTCCGTTCTTCCCTTCTTCG	120
Db	61 CCGGTACACTTGGCAGCGCCCTTAGCGCCGCTCTTCCGTTCTTCCCTTCTTCG	120
OY	121 CCAGCTTCCGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	180
Db	121 CCAGCTTCCGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	180
OY	121 CCAGCTTCCGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	180
Db	121 CCAGCTTCCGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	180
OY	181 TTATGCTTACGGCACTGACCCCAAAACTTGAATGAGGTGATGGTTACGTAGTG	240
Db	181 TTATGCTTACGGCACTGACCCCAAAACTTGAATGAGGTGATGGTTACGTAGTG	240
OY	241 GGCATATGCGGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	300
Db	241 GGCATATGCGGCGCTTCCCGCTCAAGCTCTAATCGGGGCTCCCTTAGGGTTCGAT	300
OY	301 GTGACCTCTTGTCCAAACTGGAACAACACTCAACCTTCTCGGCTCTTCTTGAAT	360
Db	301 GTGACCTCTTGTCCAAACTGGAACAACACTCAACCTTCTCGGCTCTTCTTGAAT	360
OY	361 TATTAAGGATTTTTCGATTTGGGCTATTTGGTTAAATAAGCGATTTAAACAAAT	420
Db	361 TATTAAGGATTTTTCGATTTGGGCTATTTGGTTAAATAAGCGATTTAAACAAAT	420
OY	421 TTAACGGCAATTTTAAACAAATATTAACGCTTACATTTCCATTGCGCATTCAGGCTCG	480
Db	421 TTAACGGCAATTTTAAACAAATATTAACGCTTACATTTCCATTGCGCATTCAGGCTCG	480
OY	481 CAACCTGTTGGGAAGGGGATCGGTGGGCTCTTCTGCTATTACGCAAGCTGGCAAAAG	540
Db	481 CAACCTGTTGGGAAGGGGATCGGTGGGCTCTTCTGCTATTACGCAAGCTGGCAAAAG	540
OY	541 GGGATGTGCGAAGGGGATTAAGTTGGGTAACGCCAGGGTTTCCCATCGACGAGTGG	600
Db	541 GGGATGTGCGAAGGGGATTAAGTTGGGTAACGCCAGGGTTTCCCATCGACGAGTGG	600
OY	601 TAAACAGCGCGCAGTGAAGCTGTATTAATAGTATTAATCAATTACGGGGTCATTAGT	660
Db	601 TAAACAGCGCGCAGTGAAGCTGTATTAATAGTATTAATCAATTACGGGGTCATTAGT	660
OY	661 CATAGCCCATATATGAGTTCCGCTTACATACCTTACGTAATAGCCCGCGCTGCTGA	720
Db	661 CATAGCCCATATATGAGTTCCGCTTACATACCTTACGTAATAGCCCGCGCTGCTGA	720
OY	721 CCGGCCAAGACCCCGCCCATTTGACGTCAATATACGTATGTTCCCATAGTAACGCCA	780
Db	721 CCGGCCAAGACCCCGCCCATTTGACGTCAATATACGTATGTTCCCATAGTAACGCCA	780



Db	4021	ACAAACCAACAGTGAATGACAGTGAATAAAAAATGCTTATATGTGAATTTGTGATGCTAT	4080
Qy	4081	TGCTTTATTTGTATACCATTTATTAAGCTGCAATTAACCAAGTAAACAACCAATTTGCTATCA	4140
Db	4081	TGCTTTATTTGTATTAACCATTTATTAAGCTGCAATTAACCAAGTAAACAACCAATTTGCTATCA	4140
Qy	4141	TTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTTTTTAAAGCAAGTAAACCTCTA	4200
Db	4141	TTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTTTTTAAAGCAAGTAAACCTCTA	4200
Qy	4201	CAAAATGTGTATGCTGATTAATGATTCCTGCTCCGCGCTTTCGGTATGATGACGTGAAAAC	4260
Db	4201	CAAAATGTGTATGCTGATTAATGATTCCTGCTCCGCGCTTTCGGTATGATGACGTGAAAAC	4260
Qy	4261	CTTCGACACATGACAGCTCCCGGAGAGGTCACAGCTTGTCTGTAAAGCGGATGCCGGGAGC	4320
Db	4261	CTTCGACACATGACAGCTCCCGGAGAGGTCACAGCTTGTCTGTAAAGCGGATGCCGGGAGC	4320
Qy	4321	AGACAAGCCCGTTCAGGGCGCTCAGCGGCTGTGGCGGTCGCGGGCGCAGCCATGACC	4380
Db	4321	AGACAAGCCCGTTCAGGGCGCTCAGCGGCTGTGGCGGTCGCGGGCGCAGCCATGACC	4380
Qy	4381	CAGTCACGTACCGATAGCGAGTGTATACTGTGCTTAATGCGGCATCAGAGCAGATG	4440
Db	4381	CAGTCACGTACCGATAGCGAGTGTATACTGTGCTTAATGCGGCATCAGAGCAGATG	4440
Qy	4441	TACTGAGAGTCACCATATGTCGGGCGCGCTTGTGGCGGTTTTCCATAGGCTCCGGCCC	4500
Db	4441	TACTGAGAGTCACCATATGTCGGGCGCGCTTGTGGCGGTTTTCCATAGGCTCCGGCCC	4500
Qy	4501	CCTGACGAGCATCACAAAAATGACGCTCAAGTACAGAGTGGCGAACCAGCAGAGACTA	4560
Db	4501	CCTGACGAGCATCACAAAAATGACGCTCAAGTACAGAGTGGCGAACCAGCAGAGACTA	4560
Qy	4561	TAAAGATACCAAGGGGTTTTCCCTGTGAAGCTCCTGTGGCGCTCTCCTGTTCGACCTG	4620
Db	4561	TAAAGATACCAAGGGGTTTTCCCTGTGAAGCTCCTGTGGCGCTCTCCTGTTCGACCTG	4620
Qy	4621	CCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAAGTGGCGCTTTCATATAGC	4680
Db	4621	CCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAAGTGGCGCTTTCATATAGC	4680
Qy	4681	TCACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTTCGCTCAAGCTGGGCTGTGTGCAC	4740
Db	4681	TCACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTTCGCTCAAGCTGGGCTGTGTGCAC	4740
Qy	4741	GAACCCCCCTTCAGCCCGACCGCTTATCCGGTAACTATCGTCTTGAGTCCAC	4800
Db	4741	GAACCCCCCTTCAGCCCGACCGCTTATCCGGTAACTATCGTCTTGAGTCCAC	4800
Qy	4801	CCGCTTAAAGACAGCACTTATCCGCACTGGACAGCCACTGCTTAACAGAGATTAGCAGACG	4860
Db	4801	CCGCTTAAAGACAGCACTTATCCGCACTGGACAGCCACTGCTTAACAGAGATTAGCAGACG	4860
Qy	4861	AGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCCTTAACCTACGCGCTACACTAGA	4920
Db	4861	AGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCCTTAACCTACGCGCTACACTAGA	4920
Qy	4921	AGGACAGTATTTGGTATCTGTGCGTCTCTGTAAGCAGTTACTTCGGAATAAGATGGTGT	4980
Db	4921	AGGACAGTATTTGGTATCTGTGCGTCTCTGTAAGCAGTTACTTCGGAATAAGATGGTGT	4980
Qy	4981	AGCTTGTATCCCGCAACAACAACGACGCGTGTAGCGGTGTTTTTTTGTGTGAACGAG	5040
Db	4981	AGCTTGTATCCCGCAACAACAACGACGCGTGTAGCGGTGTTTTTTTGTGTGAACGAG	5040
Qy	5041	CAGATTACGCGCAGAAAAAGATCTCAGAAAGATCCTTTGATCTTTCTACGGGGCT	5100
Db	5041	CAGATTACGCGCAGAAAAAGATCTCAGAAAGATCCTTTGATCTTTCTACGGGGCT	5100
Qy	5101	GACGCTAGTGAACGAAAACTACGTTAAGGATTTGTGCATGAGATTATCAAAAAAG	5160
Db	5101	GACGCTAGTGAACGAAAACTACGTTAAGGATTTGTGCATGAGATTATCAAAAAAG	5160
Db	5101	GACGCTAGTGAACGAAAACTACGTTAAGGATTTGTGCATGAGATTATCAAAAAAG	5160
Qy	5161	ATCTTCACCTGATACCTTTTAAATTAAGTATTAAATCAATCAATTAAGATATAT	5220
Db	5161	ATCTTCACCTGATACCTTTTAAATTAAGTATTAAATCAATCAATTAAGATATAT	5220
Qy	5221	GAGTAAACTGTGTGACAGTTTACCAATGCTTAATCAGTAGAGGCACTATCTCAGGATC	5280
Db	5221	GAGTAAACTGTGTGACAGTTTACCAATGCTTAATCAGTAGAGGCACTATCTCAGGATC	5280
Qy	5281	TGCTATTTTGTCTATCATATAGTTTGCCTGACTCCCGCTCGTGTATGATTAAGATACG	5340
Db	5281	TGCTATTTTGTCTATCATATAGTTTGCCTGACTCCCGCTCGTGTATGATTAAGATACG	5340
Qy	5341	GAGGCGTTTACCATGTGGCCCGACAGTGTGCATGATGATGAGGAGCCACGCTCAGCGCT	5400
Db	5341	GAGGCGTTTACCATGTGGCCCGACAGTGTGCATGATGATGAGGAGCCACGCTCAGCGCT	5400
Qy	5401	CCAGATTTATCAGCAATTAACCAAGCCAGCGGAAGGCGCAGCCAGAAAGTGTCTGCA	5460
Db	5401	CCAGATTTATCAGCAATTAACCAAGCCAGCGGAAGGCGCAGCCAGAAAGTGTCTGCA	5460
Qy	5461	ACTTTATCCGCTTCATACAGTCTATTAATTTGTTCCGGGAGGTAAGTAAAGTATCG	5520
Db	5461	ACTTTATCCGCTTCATACAGTCTATTAATTTGTTCCGGGAGGTAAGTAAAGTATCG	5520
Qy	5521	CCAGTTAATGTTTGGCCCAAGGTTGTGGCATTCCTACAGGACATCGTGTACGCTCG	5580
Db	5521	CCAGTTAATGTTTGGCCCAAGGTTGTGGCATTCCTACAGGACATCGTGTACGCTCG	5580
Qy	5581	TCGTTTGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGCGATTAATGATTC	5640
Db	5581	TCGTTTGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGCGATTAATGATTC	5640
Qy	5641	CCCATGTTTGTGCAAAAAAGGCTTAGCTCCTGCTGCTCCGATCGTTGTACAGATAG	5700
Db	5641	CCCATGTTTGTGCAAAAAAGGCTTAGCTCCTGCTGCTCCGATCGTTGTACAGATAG	5700
Qy	5701	TTGGCCGCAAGTATACATCATGTATGTGACACACACTGCAATATCTCTTACATGATG	5760
Db	5701	TTGGCCGCAAGTATACATCATGTATGTGACACACACTGCAATATCTCTTACATGATG	5760
Qy	5761	CCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACAAGTCAATCTGAGAAATG	5820
Db	5761	CCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACAAGTCAATCTGAGAAATG	5820
Qy	5821	TGTTATGGGGGACGAGTTGTCTTGGCCCGGCGCATTAAGGGATTAATACCGGCGCAT	5880
Db	5821	TGTTATGGGGGACGAGTTGTCTTGGCCCGGCGCATTAAGGGATTAATACCGGCGCAT	5880
Qy	5881	AGCAGAACTTTAAAGTGTATCATATTTGAAGAGTTCTTCGGGGCGAAAACTCTCAAG	5940
Db	5881	AGCAGAACTTTAAAGTGTATCATATTTGAAGAGTTCTTCGGGGCGAAAACTCTCAAG	5940
Qy	5941	ATCTTACCGCTGTTGAGATCCAGTTGATGTAACCCACTGTGCACCCAACTGATCTTCA	6000
Db	5941	ATCTTACCGCTGTTGAGATCCAGTTGATGTAACCCACTGTGCACCCAACTGATCTTCA	6000
Qy	6001	GCATCTTTTACTTTCACACAGCTTTCGGGTGAGCAAAAAACAGAAAGCAAAATGCCGA	6060
Db	6001	GCATCTTTTACTTTCACACAGCTTTCGGGTGAGCAAAAAACAGAAAGCAAAATGCCGA	6060
Qy	6061	AAAAAGGAAATTAAGGGGACACGGAATGTTGAATTAATCAATCTTCTTTTCAATAT	6120
Db	6061	AAAAAGGAAATTAAGGGGACACGGAATGTTGAATTAATCAATCTTCTTTTCAATAT	6120
Qy	6121	TATTAAGCAATTAATCAGGGTATTTGTCTATGAGCGGATACATATTTGAATGATTTAG	6180
Db	6121	TATTAAGCAATTAATCAGGGTATTTGTCTATGAGCGGATACATATTTGAATGATTTAG	6180
Qy	6181	AAAAATTAACAAATAGGGGTTCCGCGACATTTTCCCGAAAAAGTGCAC	6229
Db	6181	AAAAATTAACAAATAGGGGTTCCGCGACATTTTCCCGAAAAAGTGCAC	6229



RESULT 2  
AR067813 LOCUS AR067813 9709 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 5 from patent US 5851813.  
ACCESSION AR067813  
VERSION AR067813.1 GI:5999035  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 9709)  
AUTHORS Desrosiers,R.C.  
TITLE Primate lentivirus antigenic compositions  
JOURNAL Patent: US 5851813-A 5 22-DEC-1998;  
FEATURES  
Source Location/Qualifiers  
1..9709  
BASE COUNT 3421 a 1756 c 2366 g 2166 t  
ORIGIN  
Query Match 43.0%; Score 2676.2; DB 6; Length 9709;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1291 CCATGAGAGTGAAGAGAGATATCAGCACTGTGGAGATGGGGTGAATGGGGACCA 1350  
DB 6219 CAATGAGAGTGAAGAGAGATATCAGCACTGTGGAGATGGGGTGAATGGGGACCA 6278  
QY 1351 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAAATTTGGGTCAACCTCTATT 1410  
DB 6279 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAAATTTGGGTCAACCTCTATT 6338  
QY 1411 ATGGGGTACCTGTGTGGAGAGACACACCACCTCTATTGTCATCAGATGCTAAAG 1470  
DB 6339 ATGGGGTACCTGTGTGGAGAGACACACCACCTCTATTGTCATCAGATGCTAAAG 6398  
QY 1471 CATATGATACAGAGTACATATGTTGGGCCACACATGCGGTACCACAGACCCCA 1530  
DB 6399 CATATGATACAGAGTACATATGTTGGGCCACACATGCGGTACCACAGACCCCA 6458  
QY 1531 ACCCAAGAGATGATTTGGTAAATGTGACAGAAAAATTTAACTGTGAAAAATGACA 1590  
DB 6459 ACCCAAGAGATGATTTGGTAAATGTGACAGAAAAATTTAACTGTGAAAAATGACA 6518  
QY 1591 TGTGTAAGACAGATGATGAGATATATCACTTATGGGATCAAAAGCCTAAAGCCATGTG 1650  
DB 6519 TGTGTAAGACAGATGATGAGATATATCACTTATGGGATCAAAAGCCTAAAGCCATGTG 6578  
QY 1651 TAAATTTAACCCCACTGCTGTGTTAAAGTGCACATGTTGAAGAAATGATACTATA 1710  
DB 6579 TAAATTTAACCCCACTGCTGTGTTAAAGTGCACATGTTGAAGAAATGATACTATA 6638  
QY 1711 CCAATAGTACTAGCGGAGATGATATATGAGAAAAAGAGATAAAAAGCTCTTCA 1770  
DB 6639 CCAATAGTACTAGCGGAGATGATATATGAGAAAAAGAGATAAAAAGCTCTTCA 6698  
QY 1771 ATATCAGCACAAGCATTAAGATTAAGTGCAGAAAAAGATATGATCTTTTATAAAGCTTG 1830  
DB 6699 ATATCAGCACAAGCATTAAGATTAAGTGCAGAAAAAGATATGATCTTTTATAAAGCTTG 6758  
QY 1831 ATATAGTACCAATAGATTAATACCACTATAGTGTATAGTTGTAAACCTCAGCATTA 1890  
DB 6759 ATATAGTACCAATAGATTAATACCACTATAGTGTATAGTTGTAAACCTCAGCATTA 6818  
QY 1891 CACAGAGCTGTCCAAAGGTATCTTTGAGCAATCCATACATATTTGTGCCCCGGGTG 1950  
DB 6819 CACAGAGCTGTCCAAAGGTATCTTTGAGCAATTTCCATATTTATTTGTGCCCCGGGTG 6878  
QY 1951 GTTTTGCATCTTAAATATGATATAAAGATTCATGAAAGACCATGTAACAAATG 2010  
DB 6879 GTTTTGCATCTTAAATATGATATAAAGATTCATGAAAGACCATGTAACAAATG 6938

QY 2011 TCAGCAGATCAATATGATACATGGAATCAGGCCAGTAGTATCAACTCACTGCTTTAA 2070  
DB 6939 TCAGCAGATCAATATGATACATGGAATCAGGCCAGTAGTATCAACTCACTGCTTTAA 6998  
QY 2071 ATGGCAGTCTAGCAGAAAGATGATGATTAATGATCTGCCAATTTCCACAGATGCTA 2130  
DB 6999 ATGGCAGTCTAGCAGAAAGATGATGATTAATGATCTGCCAATTTCCACAGATGCTA 7058  
QY 2131 AAACCATATATGATCAGTGTGAACACATCTGTGAATTAATTTGTACAAAGACCAACA 2190  
DB 7059 AAACCATATATGATCAGTGTGAACACATCTGTGAATTAATTTGTACAAAGACCAACA 7118  
QY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGCATTTGTTCAATAGGAA 2250  
DB 7119 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGCATTTTTCATAGGAA 7178  
QY 2251 AAATAGAAATATGAGACACACATTTGATTAATAGAGCAAAATGGAATGCCACTT 2310  
DB 7179 AAATAGAAATATGAGACACACATTTGATTAATAGAGCAAAATGGAATGCCACTT 7238  
QY 2311 TAAACAGATATGCTAGCAATTTAGAGAAACAATTTGGAATATTAATTAATCTTTA 2370  
DB 7239 TAAACAGATATGCTAGCAATTTAGAGAAACAATTTGGAATATTAATTAATCTTTA 7298  
QY 2371 AGCAATCTCAGAGAGGGGACCGAGAAATTTGTAACGACAGTTTAAATGAGAGGGAAT 2430  
DB 7299 AGCAATCTCAGAGAGGGGACCGAGAAATTTGTAACGACAGTTTAAATGAGAGGGAAT 7358  
QY 2431 TTTTCTACTGTAATTTCAACACACACTTTTAATAGTACTGTTTAATAGTACTTGAGATA 2490  
DB 7359 TTTTCTACTGTAATTTCAACACACACTTTTAATAGTACTGTTTAATAGTACTTGAGATA 7418  
QY 2491 CTGAAGGTCATTAATTAACCTGAGAGAAATGACACATACACTCCCATGCAAGATAAAC 2550  
DB 7419 CTGAAGGTCATTAATTAACCTGAGAGAAATGACACATACACTCCCATGCAAGATAAAC 7478  
QY 2551 AATTTATTAACATGTCGAGAGAAAGTAAAGCAATGATGCCCCCTCCATCACTGGAC 2610  
DB 7479 AATTTATTAACATGTCGAGAGAAAGTAAAGCAATGATGCCCCCTCCATCACTGGAC 7538  
QY 2611 AATTTAGATGTCATCAATATTTACTGGGCTCTATTAAACAAGATGATGGTAAATACA 2670  
DB 7539 AATTTAGATGTCATCAATATTTACTGGGCTCTATTAAACAAGATGATGGTAAATACA 7598  
QY 2671 ACAATGGTCCGAGATCTTTCAGACCTGGAAGAGCGATATGAGGACAAATTTGGAGATG 2730  
DB 7599 ACAATGGTCCGAGATCTTTCAGACCTGGAAGAGCGATATGAGGACAAATTTGGAGATG 7658  
QY 2731 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2790  
DB 7659 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7718  
QY 2791 AGAGAAGATGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGAGACTTTGTTCCTTGGGT 2850  
DB 7719 AGAGAAGATGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGAGACTTTGTTCCTTGGGT 7778  
QY 2851 TCTTGGGAGCAGCAGGAAGCACTATGGCTGCAGCTCAATGACGCTGACAGGCCA 2910  
DB 7779 TCTTGGGAGCAGCAGGAAGCACTATGGCTGCAGCTCAATGACGCTGACAGGCCA 7838  
QY 2911 GACAAATTTGTCGATATAGTGCAGCAGCAGCAAAATTTGCTGAGGGCTATTGAGGGC 2970  
DB 7839 GACAAATTTGTCGATATAGTGCAGCAGCAGCAAAATTTGCTGAGGGCTATTGAGGGC 7898  
QY 2971 AACAGCATCTGTGCACTCAGAGTCTGGGCATCAACAGCTCCAGCAGAAATCTCGG 3030  
DB 7899 AACAGCATCTGTGCACTCAGAGTCTGGGCATCAACAGCTCCAGCAGAAATCTCGG 7958  
QY 3031 CTGTGAAAAGATACCTAAAGATCAACAGCTCTGAGGATTTGGGCTTCTGTGAAAAC 3090  
DB 7959 CTGTGAAAAGATACCTAAAGATCAACAGCTCTGAGGATTTGGGCTTCTGTGAAAAC 8018

3091 TCATTGGACCACTGCTGTGCTTGGATGCTAGTTGGAGTAATTAATCTCTGACACGA 3150  
|||||  
8019 TCATTGGACCACTGCTGTGCTTGGATGCTAGTTGGAGTAATTAATCTCTGACACGA 8078  
|||||  
3151 TTTGGATTAACATGACCTGGATGGAGTGGACAGAGAAATTACAAATTACCAAGCTTAA 3210  
|||||  
8079 TTTGGAAATTAACATGACCTGGATGGAGTGGACAGAGAAATTACAAATTACCAAGCTTAA 8138  
|||||  
3211 TACACTCCTTAATTAAGAAATCGCAAAACGACAGAAAGAAATGAACAATAATTATTTGG 3270  
|||||  
8139 TACACTCCTTAATTAAGAAATCGCAAAACGACAGAAAGAAATGAACAATAATTATTTGG 8198  
|||||  
3271 AATTAGATAAATGGGCAAGTTTGGAAATGTTTAAACATAAATTTGCTGTGTATA 3330  
|||||  
8199 AATTAGATAAATGGGCAAGTTTGGAAATGTTTAAACATAAATTTGCTGTGTATA 8258  
|||||  
3331 TAAATATTTTCATTAATGATAGTAGAGCTTGGTAAAGTTTAAAGATTAATTTTGTGTAC 3390  
|||||  
8259 TAAATATTTTCATTAATGATAGTAGAGCTTGGTAAAGTTTAAAGATTAATTTTGTGTAC 8318  
|||||  
3391 TTTCTATAGTAATAGATTAGGAGGATATTCACCATTTATCGTTTCAGACCCACCTCC 3450  
|||||  
8319 TTTCTATAGTAATAGATTAGGAGGATATTCACCATTTATCGTTTCAGACCCACCTCC 8378  
|||||  
3451 CAATCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAAGTGGAGAGAGACA 3510  
|||||  
8379 CAATCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAAGTGGAGAGAGACA 8438  
|||||  
3511 GAGACAGATCCATTCGATAGTACTGAACGATACCTTAGCATTTATCTGGAGCATCTGCGGA 3570  
|||||  
8439 GAGACAGATCCATTCGATAGTACTGAACGATACCTTAGCATTTATCTGGAGCATCTGCGGA 8498  
|||||  
3571 GCTGTGCTCTTACGCTACACCGCTTGAAGAGACTTACTCTGATTTGTAACGAGATTTG 3630  
|||||  
8499 GCTGTGCTCTTACGCTACACCGCTTGAAGAGACTTACTCTGATTTGTAACGAGATTTG 8558  
|||||  
3631 TGGAACTTCTGGAGCGAGGGGTGGGAAAGCCCTCAATATTGGTGAATCTCTACAGT 3690  
|||||  
8559 TGGAACTTCTGGAGCGAGGGGTGGGAAAGCCCTCAATATTGGTGAATCTCTACAGT 8618  
|||||  
3691 ATTGGAGTCAGAGACTTAAGAATAGTCTGTTAACTTGCCTCAATGCCACAGCCATTTAGCAG 3750  
|||||  
8619 ATTGGAGTCAGAGACTTAAGAATAGTCTGTTAACTTGCCTCAATGCCACAGCCATTTAGCAG 8678  
|||||  
3751 TAGCTGAGGGGACAGATAGGCTTATAGAAGTATTACAAGCAGCTTATAGAGCTTATGCGCC 3810  
|||||  
8679 TAGCTGAGGGGACAGATAGGCTTATAGAAGTATTACAAGCAGCTTATAGAGCTTATGCGCC 8738  
|||||  
3811 ACATACCTAGAAAGAAATAAGACAGGCTTGGAAAGATTTTCTTAAAGATGGGTGGCAAG 3870  
|||||  
8739 ACATACCTAGAAAGAAATAAGACAGGCTTGGAAAGATTTTCTTAAAGATGGGTGGCAAG 8798  
|||||  
3871 TGTGTCAAAAAAGTATGATGATGGATGGCTCTGTAAGGGAATAAATAGACGAGCTGAG 3930  
|||||  
8799 TGTGTCAAAAAAGTATGATGATGGATGGCTCTGTAAGGGAATAAATAGACGAGCTGAG 8858  
|||||  
3931 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGATCTAGA 3971  
|||||  
8859 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGATCTAGA 8899  
|||||

RESULT 3  
AR224437 9709 bp DNA linear PAT 26-SEP-2002  
LOCUS AR224437 Sequence 15 from patent US 6440730.  
DEFINITION AR224437  
ACCESSION AR224437  
VERSION AR224437.1 GI:23333237  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 9709)  
AUTHORS Von Laer, M.-D. and Beyer, W.

TITLE Retroviral hybrid vectors pseudotyped with LCMV  
JOURNAL Patent: US 6440730-A 15 27-AUG-2002;  
FEATURES location/Qualifiers  
SOURCE 1. 9709  
/organism="unknown"  
BASE COUNT 3421 a 1756 c 2366 g 2166 t  
ORIGIN  
Query Match 43.0%; Score 2676.2; DB 6; Length 9709;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1291 CCATGAGAGTGAAGAGAGATATCAGCAGCTTGTGAGATGGGGTGGAAATGGGGACCA 1350  
|||||  
6219 CAATGAGAGTGAAGAGAGATATCAGCAGCTTGTGAGATGGGGTGGAAATGGGGACCA 6278  
|||||  
1351 TGTCTCTTGGGATTTGTATGATCTGTATGCTGTACAGAAATTTGGGTACCTCTATT 1410  
|||||  
6279 TGTCTCTTGGGATTTGTATGATCTGTATGCTGTACAGAAATTTGGGTACCTCTATT 6338  
|||||  
1411 ATGGGGTACCTGTGGAGAGAGCAACACACCTCTATTTTGGATCAGATGCTTAAG 1470  
|||||  
6339 ATGGGGTACCTGTGGAGAGAGCAACACACCTCTATTTTGGATCAGATGCTTAAG 6398  
|||||  
1471 CATATGATACAGAGTACATATGTTTGGGACACATGCTGTGTACCCACAGACCCA 1530  
|||||  
6399 CATATGATACAGAGTACATATGTTTGGGACACATGCTGTGTACCCACAGACCCA 6458  
|||||  
1531 ACCCAAGAAGTATGTTTGGTAAATGTGACAGAAATTTTACATGTGAAAAATACA 1590  
|||||  
6459 ACCCAAGAAGTATGTTTGGTAAATGTGACAGAAATTTTACATGTGAAAAATACA 6518  
|||||  
1591 TGTAGAGACAGATCATGAGAGATTAATCAAGTTTGGATGCAAGCCTTAAGCATGTG 1650  
|||||  
6519 TGTAGAGACAGATCATGAGAGATTAATCAAGTTTGGATGCAAGCCTTAAGCATGTG 6578  
|||||  
1651 TAAATTAACCCACTCTGTGTATGTTTAAAGTCACTGATTTGAGACATGATATATA 1710  
|||||  
6579 TAAATTAACCCACTCTGTGTATGTTTAAAGTCACTGATTTGAGACATGATATATA 6638  
|||||  
1711 CCATATGATAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAATCGCTCTTCA 1770  
|||||  
6639 CCATATGATAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAATCGCTCTTCA 6698  
|||||  
1771 ATATCAGACAAAGCATAGAGATTAAGTGAAGAAAGATTAATGCAATCTTTTAAACTTG 1830  
|||||  
6699 ATATCAGACAAAGCATAGAGATTAAGTGAAGAAAGATTAATGCAATCTTTTAAACTTG 6758  
|||||  
1831 ATATGATACCAATAGATTAATCAAGCTATAGTGTATAGTGTATACACCTCACTCTTA 1890  
|||||  
6759 ATATGATACCAATAGATTAATCAAGCTATAGTGTATAGTGTATACACCTCACTCTTA 6818  
|||||  
1891 CACAGGCTGTCCAAAGGATATCCTTTGAGCAATTTCCCATATATTTGTCCTGCTG 1950  
|||||  
6819 CACAGGCTGTCCAAAGGATATCCTTTGAGCAATTTCCCATATATTTGTCCTGCTG 6878  
|||||  
1951 GTTTTGGATTTTAAATGTAATTAAGAAGTTCAATGGAACAGACATGATACAATG 2010  
|||||  
6879 GTTTTGGATTTTAAATGTAATTAAGAAGTTCAATGGAACAGACATGATACAATG 6938  
|||||  
2011 TCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2070  
|||||  
6939 TCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6998  
|||||  
2071 ATGGCAGCTAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130  
|||||  
6999 ATGGCAGCTAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7058  
|||||  
2131 AAACCATTAATAGTACAGCTGATACACATCTGTAGAAATTAATGATACAAAGCCCAACACA 2190  
|||||  
7059 AAACCATTAATAGTACAGCTGATACACATCTGTAGAAATTAATGATACAAAGCCCAACACA 7118  
|||||  
2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGACATTTTGTATACATAGGAA 2250  
|||||

```

Db      7119 ATACAGAAAAAGTCCGTATCCAGAGGGGACAGGAGACATTTGTACAAATAGAA 7178
Qy      2251 AAATGGAATATGAGACACATTTGTAACATTAAGACAAATAGATGACACTT 2310
Db      7179 AAATGGAATATGAGACACATTTGTAACATTAAGACAAATAGATGACACTT 7238
Qy      2311 TAAACAGATAGCTAGCAATTAAGAGACATTTGGAATTAATAAACAATTAATCTTTA 2370
Db      7239 TAAACAGATAGCTAGCAATTAAGAGACATTTGGAATTAATAAACAATTAATCTTTA 7298
Qy      2371 AGCAATCTCAGAGAGGACCCAGAAATTTGAACGACATTTTAATTTGAGGGGAAT 2430
Db      7299 AGCAATCTCAGAGAGGACCCAGAAATTTGAACGACATTTTAATTTGAGGGGAAT 7358
Qy      2431 TTTTCTACTGTAATCAACACACAGTTTAATAGACTGTTTAATAGTACTTGAGATA 2490
Db      7359 TTTTCTACTGTAATCAACACACAGTTTAATAGACTGTTTAATAGTACTTGAGATA 7418
Qy      2491 CTGAAGGGTCAAAATTAACACTGAGAGAGTGAACACATCCATCCATGCAATTAAC 2550
Db      7419 CTGAAGGGTCAAAATTAACACTGAGAGAGTGAACACATCCATCCATGCAATTAAC 7478
Qy      2551 AATTTATTAACATGTGGCAGAGAGTGAAGAAAAGCAATGTATGCCCTCCATCAGTGAC 2610
Db      7479 AATTTATTAACATGTGGCAGAGAGTGAAGAAAAGCAATGTATGCCCTCCATCAGTGAC 7538
Qy      2611 AAATAGATGTTTCATCAAAATTAATGCTGGCTGCTTTTAACAAGATGCTGTAATACA 2670
Db      7539 AAATAGATGTTTCATCAAAATTAATGCTGGCTGCTTTTAACAAGATGCTGTAATACA 7598
Qy      2671 ACAATGGTCCGAGATCTTACAGACTGAGAGAGGAGATGAGAGCAATTTGAGAAAGTG 2730
Db      7599 ACAATGGTCCGAGATCTTACAGACTGAGAGAGGAGATGAGAGCAATTTGAGAAAGTG 7658
Qy      2731 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2790
Db      7659 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7718
Qy      2791 AGAGAGAGTGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2850
Db      7719 AGAGAGAGTGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7778
Qy      2851 TCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2910
Db      7779 TCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7838
Qy      2911 GACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2970
Db      7839 GACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7898
Qy      2971 AACAGCATCTGTGCACTCAGAGTCTGGGAGATTAACAGCTCCAGCAAGATCTGGG 3030
Db      7899 AACAGCATCTGTGCACTCAGAGTCTGGGAGATTAACAGCTCCAGCAAGATCTGGG 7958
Qy      3031 CTGTGGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3090
Db      7959 CTGTGGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8018
Qy      3091 TCATTGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3150
Db      8019 TCATTGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8078
Qy      3151 TTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3210
Db      8079 TTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8138
Qy      3211 TACAATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3270
Db      8139 TACAATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8198
Qy      3271 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3330

```

```

Db      8199 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8258
Qy      3331 TAAATTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3390
Db      8259 TAAATTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8318
Qy      3391 TTTCTATAGTGAATAGATTAAGCAGGATATTCACATTAATTCGTTTACAGACCCACCTCC 3450
Db      8319 TTTCTATAGTGAATAGATTAAGCAGGATATTCACATTAATTCGTTTACAGACCCACCTCC 8378
Qy      3451 CAATCCGAGGGGACCCGACAGGCCGGAAGAAATGAAGAAGAAAGAGAGAGAGACA 3510
Db      8379 CAATCCGAGGGGACCCGACAGGCCGGAAGAAATGAAGAAGAAAGAGAGAGAGACA 8438
Qy      3511 GAGACAGATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3570
Db      8439 GAGACAGATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8498
Qy      3571 GCCTGTGCTCTTACAGTACACACGCTTGAGAGACTTACTCTGTAATTAAGAGATTTG 3630
Db      8499 GCCTGTGCTCTTACAGTACACACGCTTGAGAGACTTACTCTGTAATTAAGAGATTTG 8558
Qy      3631 TGGAACTTCTGGGACGACAGGGGGTGGGAAGCCCTCAATTAATTTGGTAATCTCTACAGT 3690
Db      8559 TGGAACTTCTGGGACGACAGGGGGTGGGAAGCCCTCAATTAATTTGGTAATCTCTACAGT 8618
Qy      3691 ATTGAGTCCAGAACTTAAGAAATAGTCTGTTAACTTCTCAATTCACACAGCCATAGAG 3750
Db      8619 ATTGAGTCCAGAACTTAAGAAATAGTCTGTTAACTTCTCAATTCACACAGCCATAGAG 8678
Qy      3751 TAGCTGAGGGGACAGATAGGTTATGAAGATTAATTAACAGACCTTAATAGATTAATTCGC 3810
Db      8679 TAGCTGAGGGGACAGATAGGTTATGAAGATTAATTAACAGACCTTAATAGATTAATTCGC 8738
Qy      3811 ACATACCTAGAAGAAATTAAGACAGGGCTTGAAGAGATTTGCTATTAATAGATGGTGGCAAG 3870
Db      8739 ACATACCTAGAAGAAATTAAGACAGGGCTTGAAGAGATTTGCTATTAATAGATGGTGGCAAG 8798
Qy      3871 TGGTCAAAAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3930
Db      8799 TGGTCAAAAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8858
Qy      3931 CCAGCAGCAGATGGGGTGGGAGACAGTATCTCGAGATCTAGA 3971
Db      8859 CCAGCAGCAGATGGGGTGGGAGACAGTATCTCGAGACCTAGA 8899

```

```

RESULT 4
AX032749 9709 bp DNA Linear PAT 21-SEP-2000
LOCUS
DEFINITION Sequence 1 from Patent W00047223.
ACCESSION AX032749
VERSION AX032749.1 GI:10279727
KEYWORDS
SOURCE
ORGANISM Human immunodeficiency virus
            Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            Lentivirus group.
REFERENCE
  1. Schreiber, M.
  2. Viral vaccine
  3. Patent: WO 0047223-A 1 17-AUG-2000;
  4. SCHREIBER MICHAEL (DE) ; STRATHMAN AG & CO (DE)
FEATURES
  1. source
  2. /organism="Human immunodeficiency virus"
  3. /mol_type="genomic DNA"
  4. /db_xref="taxon:12721"
BASE COUNT 3421 a 1756 c 2366 g 2166 t
ORIGIN
Query Match 43.0%; Score 2676.2; DB 6; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;

```



```

Oy 3451 CAATCCGAGGGGACCCGACAGGCCGGAAGGATGAAGAAGAGTGGAGAGACA 3510
Db 8379 CAATCCGAGGGGACCCGACAGGCCGGAAGGATGAAGAAGAGTGGAGAGACA 8438
Oy 3511 GAGACAGATCCATTGATAGTGAACGGATCTTACGACTTATCTGGAGCATCTGGCGA 3570
Db 8439 GAGACAGATCCATTGATAGTGAACGGATCTTACGACTTATCTGGAGCATCTGGCGA 8498
Oy 3571 GCGTGTGCTCTTACGATACCGCTTGGAGAGCTTACTTGAATGTAAGAGATGG 3630
Db 8499 GCGTGTGCTCTTACGATACCGCTTGGAGAGCTTACTTGAATGTAAGAGATGG 8558
Oy 3631 TGAAGCTTCTGGAGAGGGGGTGGAGGCCCTCAATATTTGGTGAATCTCTACAGT 3690
Db 8559 TGAAGCTTCTGGAGAGGGGGTGGAGGCCCTCAATATTTGGTGAATCTCTACAGT 8618
Oy 3691 ATTGAGACTCAGAACTAAGAAATAGTGTCTTAACTTCTCAATGCGACAGCCATAGCAG 3750
Db 8619 ATTGAGACTCAGAACTAAGAAATAGTGTCTTAACTTCTCAATGCGACAGCCATAGCAG 8678
Oy 3751 TAGCTGAGGGAGCAGATAGGGTTATAGAGTATTACAAGCAGCTTATAGAGCTATTCGCC 3810
Db 8679 TAGCTGAGGGAGCAGATAGGGTTATAGAGTATTACAAGCAGCTTATAGAGCTATTCGCC 8738
Oy 3811 ACATACCTAGAAAGATAAGACAGGGCTTGGAAAGGATTTGCTATAGATGGTGGCAAG 3870
Db 8739 ACATACCTAGAAAGATAAGACAGGGCTTGGAAAGGATTTGCTATAGATGGTGGCAAG 8798
Oy 3871 TGTGTCAAAAAGTACTGTGATTTGATGCTCTCTTAAGGAAAGATAGAGCAGCTGAG 3930
Db 8799 TGTGTCAAAAAGTACTGTGATTTGATGCTCTCTTAAGGAAAGATAGAGCAGCTGAG 8858
Oy 3931 CCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGACTAG 3971
Db 8859 CCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGACTAG 8899

```

## RESULT 5

```

LOCUS HIVNL43 9709 bp ss-RNA linear VRL 02-AUG-1993
DEFINITION Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant
clone pNL4-3.
VERSION M19921
KEYWORDS M19921.1 GI:328415
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.

```

```

REFERENCE 1 (bases 1 to 9709)
AUTHORS Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R.,
TITLE Radson,A. and Martin,M.A.
JOURNAL Production of acquired immunodeficiency syndrome-associated
MEDLINE retrovirus in human and nonhuman cells transfected with an
        86281827 infectious molecular clone
        J. Virol. 59 (2), 284-291 (1986)
REFERENCE 2 (bases 1 to 9709)
AUTHORS Buckler,C.E., Buckler-White,A.J., Willey,R.L. and McCoy,J.
TITLE Unpublished (1988)
REFERENCE 3 (sites)
AUTHORS Dai,L.C., West,K., Littau,R., Takahashi,K. and Ennis,F.A.
TITLE Mutation of human immunodeficiency virus type 1 at amino acid 585
        on gp41 results in loss of killing by CD8+ A24-restricted cytotoxic
        T lymphocytes
        J. Virol. 66 (5), 3151-3154 (1992)
JOURNAL 92219406
MEDLINE 1373204
COMMENT

```

```

Original source text: Human immunodeficiency virus type 1 (HIV-1),
NY5/BRU (LAV-1) recombinant clone pNL4-3.
[3] sites; revisions of [3].
Clean copy of sequence [3] kindly provided by Chuck Buckler, NIAID,

```

Bethesda, MD, 24-JUN-1988. The construction of pNL4-3 has been described in [1]. pNL4-3 is a recombinant (infectious) proviral clone that contains DNA from HIV isolates NY5 (5' half) and BRU (3' half). The site of recombination is the EcoRI site at positions 5743-5748.

The length and sequence of the vpr coding region corresponds to that of the BRU, SC, SP2, MAL and ELI isolates. The vpr coding region of these isolates is about 18 amino acid residues longer than the vpr coding region of the IIB isolates. In HIVNL43, this shift is due to a single base deletion (with respect to the IIB's) at position 5770. The sequence at this position is 'atttc' in HIVNL43 and 'atttc' in HIVXB2.

The original BRU clone, sequenced by Martin-Hobson, et al. (Cell 40, 9-17 (1985)), and the BRU portion of the pNL4-3 recombinant clone are different clones reported in the same BRU isolate.

Two of the revisions reported in the FEATURES produced changes in amino acid sequences. The revision at position 2421 changes one amino acid residue from 'R' to 'G' in the pol coding region. The revision at positions 8995-9000 changes three amino acid residues from 'AHT' to 'VTP' in the nef coding region.

## FEATURES

```

source          Location/Qualifiers
                1..9709
                /organism="Human immunodeficiency virus 1"
                /mol_type="genomic RNA"
                /db_xref="taxon:11676"
                1..634
                /note="5' LTR"
                repeat_region
                454..550
                /note="R repeat 5' copy"
                455..9626
                /note="tat, rev, nef subgenomic mRNA"
                744..5776
                /note="tat, rev, nef mRNA intron 1"
                790..2292
                /note="gag polyprotein"
                /codon_start=1
                /protein_id="AA44987.1"
                /db_xref="GI:328418"
                /translation="WGARASVLSGGLDKEKIRLRPGCKQYKILIIIVASRELEF
                AVNPGLETSEGRKQIIGLOLPDSLQTSSEELRSLYNTIAYLCVHORIDVKTREAD
                KIEEONKRSKKAQQAADGNNNSOVNPYVONLOGOVHQAISPTYNAMWYVE
                EKARPEVIMPESALSSEGAPODINTMLNTVGHQAMOMLKETINEEAEMDLHYE
                HAGIAPGOMREPRGSDIAGTSTLQDGMTHNPILPGEIKRYIIIGLKTIVM
                XSPSTIIDIRGKPREPRDVDPRTYLRADQASQYKMMKMTFIIYONNGLKTVK
                KAUGGATTEBMTACOGVGPGRKAVYLAEMASQYTNPATIMOKNFRNOKRTVC
                FNCGEKHLAKNCAPRRKGCWCKGEGHOMKDETQANFLGIMPSHGKRGPNFLQ
                SRPEPTAPPESEFGEETTPSOKOPEIDKELYPILASLSLFGSDPSQ"
                2085..5096
                /partial
                /note="pol polyprotein (NH2-terminus uncertain)"
                /codon_start=1
                /protein_id="AA44988.1"
                /db_xref="GI:328419"
                /translation="FFREDLAFPGKAREFSEQRTANSPRTRELQVGRDNNLSSEA
                GADRGVTSFSPQITLMORPIVTKIGGOLKEALDPTGADDTVLEEMNPGRRPKM
                IGCTGGRIVKYNQDOLILEICGHAKICTVAGPPVNVILGINTLOIGTCLNIPISPI
                ETYVYKLRPGADGPKVQMPLETERKIALVETCTEMKESGKISITGEPNPNTPYRI
                KKQSTWRKLVDFRELNRKTRQDWEVQVLIIPRAGLKQKSVTVLDVGAVESVPID
                KDFKRYAFTPIINNETPGIRYOYVNLPGQWKSPIIFQCSMTKILEPRKQNPDIY
                IYOYMDLYQSDLEIGONRTKILELQNHLLRMFTPKDHQKEPFLIMWGYELHD
                KMTVQPIVLPKEDSMVNDIOLKVLNMAISOIYAGIKVHOLCKLGTALTEVPL
                TEAEFLAKREILKREPVQYVYDPSKDLIAEIQKQSGCOMPTQIYOEPKNTKTK
                YARKGHTNDVOKULTEAVOKIATESIVMGKTPKELPTOKETWEMWMTYQATVI
                PEMEFTVTPVLVKIMYOLEKEPIIGATFVVDGANETLKGAGYVTDGRQKVVPL
                TDTNOKTELDAILALQDSGLAEVNIYTDQYALGIIQADPDKSESLVSIIEOLIK
                KEKYVLAMPANHAKIGENQVDGLVSGIRKRVLELDICIDKAEHEKYLHSMRAMAD
                FNPBPVAKETIVASCDCOLKGEAMHGQVQSPQIMQDCTCHLEGKYLIVAVHASY
                IEAEVPAETGOETAPVPLKLAGKMPKVTYHTQNSMFTSTTVYAAACGMVAKOEFQ
                PYNQSGVYIESMKELKTIIGVQDAAEHKTRAYQNAVYIHNFRKRGKIGYSGR
                IVDLIADIQTKELQKQITKIQNFRVYIRSDRQVDMGPAKLLKMGGAAYVLDNSDI
                KVPVRRAKIIRDGKQMGDDCVASRODD"
                5041..5619
                /note="vif protein"

```

[illegible]



OY	2191	ATACAGAAAAAGTATCCGTA	TCCAGAGGGGACAGGGAGCATTTGTACAA	TAGGAA	2250	
Db	7119	ATACAGAAAAAGATCCGTA	TCCAGAGGGGACAGGGAGCATTTGTACAA	TAGGAA	7178	
OY	2251	AAATAGGAAATATGAGCAAC	CAATGTATACATTTAGAGGCAAAATGGAATCCCTT	2310		
Db	7179	AAATAGGAAATATGAGCAAC	CAATGTATACATTTAGAGGCAAAATGGAATCCCTT	7238		
OY	2311	TAAACAGATAGCTAGC	AAATTAAGAGAACAAATTTGGAAATATAAACAATTAATCTTTA	2370		
Db	7229	TAAACAGATAGCTAGC	AAATTAAGAGAACAAATTTGGAAATATAAACAATTAATCTTTA	7298		
OY	2371	AGCAATCTCAGAGAGGGAC	CCAGAAATTTGACGCACATTTTAATTTGAGGGGAA	2430		
Db	7299	AGCAATCTCAGAGAGGGAC	CCAGAAATTTGACGCACATTTTAATTTGAGGGGAA	7358		
OY	2431	TTTCTCTAGTAATTC	CAACACATGTTTAATTA	TACTTGTTTAATAGTACTTGGAGTA	2490	
Db	7359	TTTCTCTAGTAATTC	CAACACATGTTTAATTA	TACTTGTTTAATAGTACTTGGAGTA	7418	
OY	2491	CTGAAGGGTCAAAATAC	ACTGAAAGAGTGACAAATCACACTCCCATCGAGATATAAC	2550		
Db	7419	CTGAAGGGTCAAAATAC	ACTGAAAGAGTGACAAATCACACTCCCATCGAGATATAAC	7478		
OY	2551	AATTTATTAACATGTGG	CAGGAAGTAGGAAAGCAATGTATGCCCTCCATCAGTGGAC	2610		
Db	7479	AATTTATTAACATGTGG	CAGGAAGTAGGAAAGCAATGTATGCCCTCCATCAGTGGAC	7538		
OY	2611	AAATTAAGTGTCAATCA	AAATTTACTGCGCTCATTTAACAAAGATGTGTGTAATACA	2670		
Db	7539	AAATTAAGTGTCAATCA	AAATTTACTGCGCTCATTTAACAAAGATGTGTGTAATACA	7598		
OY	2671	ACAATGGGTCCGAGATCT	TACAGACCTGAGAGGCGATATAGAGGACAAATTTGGAAGTG	2730		
Db	7599	ACAATGGGTCCGAGATCT	TACAGACCTGAGAGGCGATATAGAGGACAAATTTGGAAGTG	7658		
OY	2731	AATTTATTAATTAATAA	AGTACAAAATTTGACCAATTAGAGTAGACCCACCAAGGCAA	2790		
Db	7659	AATTTATTAATTAATAA	AGTACAAAATTTGACCAATTAGAGTAGACCCACCAAGGCAA	7718		
OY	2791	AGACAGAGTGTGTGAC	AGAGAAAAAGACAGTGGGAATTTGAGACTTTGTTCTTGGGT	2850		
Db	7719	AGACAGAGTGTGTGAC	AGAGAAAAAGACAGTGGGAATTTGAGACTTTGTTCTTGGGT	7778		
OY	2851	TCTTGGGAGCAGCAG	GAAGCACTATGGGCTTCACAGTCAATGACGCTGACGGTACAGGCCA	2910		
Db	7779	TCTTGGGAGCAGCAG	GAAGCACTATGGGCTTCACAGTCAATGACGCTGACGGTACAGGCCA	7838		
OY	2911	GACAAATTAATTTGCT	GTGATATAGTGCAGCAGCAACAATTTGCTGAGGGCTTAATGAGGCGC	2970		
Db	7839	GACAAATTAATTTGCT	GTGATATAGTGCAGCAGCAACAATTTGCTGAGGGCTTAATGAGGCGC	7898		
OY	2971	AACAGCATCTTTGCCA	CTCACACTCTGCGGCATCAAAACAGCTCAGGCAAGATCTCTGG	3030		
Db	7899	AACAGCATCTTTGCCA	CTCACACTCTGCGGCATCAAAACAGCTCAGGCAAGATCTCTGG	7958		
OY	3031	CTGTGGAAGAATACCT	TAAAGGATCAACAGCTCTGGGGATTTGGGGTTGCTCTGGA	3090		
Db	7959	CTGTGGAAGAATACCT	TAAAGGATCAACAGCTCTGGGGATTTGGGGTTGCTCTGGA	8018		
OY	3091	TCATTTGCACAC	CTGCTGTGCTGCTGGAATGCTAGTTGGAGTAATTAATCTCTG	3150		
Db	8019	TCATTTGCACAC	CTGCTGTGCTGCTGGAATGCTAGTTGGAGTAATTAATCTCTG	8078		
OY	3151	TTTGGAAATACATGAC	CTCGATGGAATGGGACAAGAAATTTMACAAATTTACCAAGCTTAA	8138		
Db	8079	TTTGGAAATACATGAC	CTCGATGGAATGGGACAAGAAATTTMACAAATTTACCAAGCTTAA	8198		
OY	3211	TACACTCCTTAATTA	GGAATTCGCAAAACCGACAAAGAAATATGAACAAATTTATTTGG	3270		
Db	8139	TACACTCCTTAATTA	GGAATTCGCAAAACCGACAAAGAAATATGAACAAATTTATTTGG	8198		
OY	3271	AATTAGATTAATGGG	CAAGTTTGTGGAATTTGTTTAACATTAACAAATTTGGCTGTGTA	3330		

Db	AF324493	14824 bp	DNA	circular SYN 15-FEB-2001
Db	AF324493	14824 bp	DNA	circular SYN 15-FEB-2001
LOCUS	AF324493	14824 bp	DNA	circular SYN 15-FEB-2001
DEFINITION	HIV-1 vector pNL4-3, complete sequence.			
ACCESSION	AF324493			
VERSION	AF324493.1	GI:12831134		
KEYWORDS				
SOURCE	HIV-1 vector pNL4-3			
ORGANISM	HIV-1 vector pNL4-3			
REFERENCE	1 (bases 1 to 9709)			
AUTHORS	Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R.,			
	Rabson, A., and Martin, M. A.			
	Production of acquired immunodeficiency syndrome-associated			
	retrovirus in human and nonhuman cells transfected with an			
	infectious molecular clone			
	J. Virol. 59 (2), 284-291 (1986)			
JOURNAL	86281827			
MEDLINE	3016298			
PUBMED	2 (bases 1 to 14824)			
REFERENCE	Bosche, W. J., Poon, D. T. K., Ott, D. E., Hu, W. S. and Gorelick, R. J.			
AUTHORS	Complete Plasmid Sequence of pNL4-3			
TITLE	Unpublished			
JOURNAL	3 (bases 1 to 14824)			
REFERENCE	Bosche, W. J., Poon, D. T. K., Ott, D. E., Hu, W. S. and Gorelick, R. J.			
AUTHORS				

TITLE		Direct Submission
JOURNAL		Submitted (28-NOV-2000) AIDS Vaccine Program, SAIC Frederick and Drug Resistance Program, National Cancer Institute at Frederick, PO Box B, Frederick, Maryland 21702-1202, USA
FEATURES	source	Location/Qualifiers
	source	1. .14824
source	1. .9709	/organism="HIV-1 vector pNL4-3"
	1. .9709	/mol_type="other DNA"
source	1. .9709	/db_xref="taxon:151458"
	1. .9709	/plasmid="pNL4-3"
source	1. .9709	/focus
	1. .9709	/organism="Human immunodeficiency virus 1"
source	1. .9709	/proviral
	1. .9709	/mol_type="other DNA"
source	1. .9709	/strain="NY5/BRU (LAV-1)"
	1. .9709	/db_xref="taxon:11676"
source	1. .9709	/note="from sequence deposited in Genbank Accession Number M19921"
	1. .9709	1. .634
LTR	1. .634	/note="5' long terminal repeat"
	1. .634	454. .550
repeat_region	454. .550	/note="R: 5' copy"
	790. .2292	790. .2292
CDS	790. .2292	/codon_start=1
	790. .2292	/transl_table=11
CDS	790. .2292	/product="gag polyprotein"
	790. .2292	/protein_id="AAK08483.1"
CDS	790. .2292	/db_xref="GI:12831136"
	790. .2292	/translation="MGARASVLSGGELDKMKETIRPGCKQYIKIKIHWASRLERF AVNPGILLETSEGRQILGQLOPQSQTSEELSLNTIAYLYCVHQRIDVKTLEAD KIEEONKSKRKAQAADGNNNSQVSNPVIYONLOGOMVHOAISPTLANAVVE EKAFSEVIMPFMSALSEGATPODLNMTLNTGVGHOAMOMLKETINEEAEMWRLHY HAGIAGCOMREPRGSDIAGTSTLOEIGMMTHNPPIVGEIYKRWIILGLNKIVRY YSPSTIIDRGGPEPRDYDRFYKTLRAFOASOEYKNMTELLLOKNFRQRTQVTC KALGPGATLEEMTNACOGVCGGPKAKAYLAEMASQVYNPATIMLOKNFRQRTQVTC PNCKEGHIAKNCRAPRKGCGKCGSHOKDKTERQANFLGKIWPSHGKPGENFLQ SRPEPTAPPESEFRFGEETTPSQOKPEIDELYPVLSRLSFLGSDPSO"
CDS	790. .2292	<2085. .5096
	790. .2292	/codon_start=1
CDS	790. .2292	/transl_table=11
	790. .2292	/product="pol polyprotein"
CDS	790. .2292	/protein_id="AAK08484.1"
	790. .2292	/db_xref="GI:12831137"
CDS	790. .2292	/translation="FFREDLAPPGKARFSESEQTRANSPTRRELOVWGRDNNSLSEA GABRGSTFESFPOLITMORPLATYKIGGOLKEALDGDVTEENMLGRRKPKM IGGIGFIVROYDOLLEIGGKAIGTVLPYNNIIGNLLTOIGCTIIMPISPI EYVPAKFGMGDPKVPKQPLTLEEKIKALVETCTMEBEGKISKIGPENPNTPEVAL KKDSTVRKMRKLVDEFRLNRTQDFEWDVGLIPRAGLKOKRSVYLDGADVFEVPLD KQFRKTAFTIPISSINNETPGIRYOYNAVLPQGWKSGPAIFQCSMTKLEPPLMGYELHP IYOYMDLVGSDLEIGHRTKIEELRQHLRMGSTTTPDKRHKOKPEPLMGYELHP KMTVQGVIVPEKDSMTVDVIOKLVGLMMASQIVAGICRQOLKLLRGTKALTVEVPL TTEAELIENREILKEPVHYGVDPKSLDLEIOLKOGOCQVQIYQPEPKNLKTOK YAMKCAHNDVNDKOLTEAVOKIATISYIYMGTPFKLPIDQETWEAMWETIYQATNL PEEHEVNTPLVPLVWQLEKEPIIGAEFFYVDGANNRTKLGKAGYVVDRRQVYPL TDTNENFTLQALIHIALDQSGLEVINIVDSOVALGIIOAQPKSSELSVQIIISLLK KEVYIYAVPAHKGIGGNEQVDGLVAGIRKYLFLDGIDKAQEEHEKHSWARMASD FNLPRVAKIEIVASCDKQOLKEAMHGOVDSPTGMOLDCTHLEGVTLVAVHVSQY IAEVYIARTEGOETAYFILLKLAGRMPKRVTHHDNGSNFTSTTVKAAACMGAIKQEPGI PYRPGQGVYIESMKNLEKTIICQVARDQAEHKTAYQAAVFTIHNFRKAGVIGDSGER IVDIATDIQTELOKQITRKIQNFYVYRDSDDPWKGPAPLWKGEGAVVIGDNSDI KVVPRKRAKIIIRDYQKMGAGDCAVSRODED"
	790. .2292	2421
misc_difference	2421	/note="compared to sequence deposited in Genbank Accession Number M19921; results in arginine to glycine substitution"
	2421	/replace="g"
CDS	5041. .5619	/codon_start=1
	5041. .5619	/transl_table=11
CDS	5041. .5619	/product="vif protein"
	5041. .5619	/protein_id="AAK08482.1"
CDS	5041. .5619	/db_xref="GI:12831135"
	5041. .5619	

CDS	5041. .5619	/translation="MERNKQVMIYQVDRMKNITWKKLYKHMTISKAKDWYRHHI ESTNPKISEVHIPIGDAKLVIYTWGHTREKMHMGQVSIEMRKRRSTVDVDPD LAOLIHLYHFDCESSAIRNLTIGRTYSPREYDAGHNKVSLOYLAALAIKPKQIK PPLPSVRKLTEDRMKRPQRTKGHRGSHTMNGH"
	5041. .5619	5559. .5849
CDS	5559. .5849	/codon_start=1
	5559. .5849	/transl_table=11
CDS	5559. .5849	/product="vpr protein"
	5559. .5849	/protein_id="AAK08485.1"
CDS	5559. .5849	/db_xref="GI:12831138"
	5559. .5849	/translation="MEVAPDEGQRPQRPYNEWTLEELLESKSAVHFRPIWTANLQO HITEETDITWAGDEATIRLQDLFTFRIGCHRSRIGTVORARNGASRS"
CDS	5559. .5849	5743. .5748
	5559. .5849	/note="EcoRI site of recombination"
misc_feature	5743. .5748	join(5830. .6044,8369. .8414)
	5743. .5748	/codon_start=1
CDS	5743. .5748	/transl_table=11
	5743. .5748	/product="tat protein"
CDS	5743. .5748	/protein_id="AAK08486.1"
	5743. .5748	/db_xref="GI:12831139"
CDS	5743. .5748	/translation="MEVPDRLEPMKHHGSOPTACTNCYCKCFHCQYCFMTKALG ISYGRKRRORRRAHONSOFHOASISQPTSQSRGDPGPK"
	5743. .5748	join(5969. .6044,8369. .8643)
CDS	5969. .6044	/codon_start=1
	5969. .6044	/transl_table=11
CDS	5969. .6044	/product="rev protein"
	5969. .6044	/protein_id="AAK08487.1"
CDS	5969. .6044	/db_xref="GI:12831140"
	5969. .6044	/translation="MAGRSQSDSEELRTYVRLIKLYOSNPPNPPESTRQARRRRR WERORHISIERILSTYLGKRSHEPVLDPPLERLTLDNEDCGTSGTGVSPOI LVESPVLESGTE"
CDS	5969. .6044	6061. .6306
	5969. .6044	/codon_start=1
CDS	5969. .6044	/transl_table=11
	5969. .6044	/product="vpu protein"
CDS	5969. .6044	/protein_id="AAK08488.1"
	5969. .6044	/db_xref="GI:12831141"
CDS	5969. .6044	/translation="MOPIVAIVAAVVAIIIAIVVWSIVILEYRKILRQKIDRLDR LIBRAEDSGNESEGEVSAIVEMGEVGMHAPIDDL"
	5969. .6044	6221. .8785
CDS	6221. .8785	/codon_start=1
	6221. .8785	/transl_table=11
CDS	6221. .8785	/product="envelope polyprotein"
	6221. .8785	/protein_id="AAK08489.1"
CDS	6221. .8785	/db_xref="GI:12831142"
	6221. .8785	/translation="MRYKEKYOHLMRMGKMGKMLGILMIGSATEKLTAVTYGVYV WKATYTLRQASAKADPDEVHVMYMAHACVPDPDNPQEVLYVNTENEMKENDVRE OMHEDITSLMDQSLKPCVKLTPLCYSLKCTDKLNDNTNNSSGRMIMKEGELKCSFN ISTSIRKQKERAFFYKLDIVPIDNTSYRLISCNSTVITQACPKVSFEPIPIHGCAP AGAAILKCNKRTNGTGPCTNVSTVQCTHGIQVSPQQLNLNSLAEEDVIRISANT DNKATIIQVQNTSVEINCTPNNNTKRSIQRGPAFYTGKIGMROQACNISRA KMNATLQWTSKILREOQGNKTTIIFKQSGGDEIYHFNCGGEFFYCNSTOLFNSY WPNSTWSTBESNNGSDTTLTCRIKOFINMNOGCKAMVADPISGOJCSNITGL LTRDGGNNNNGSEIIRPGGDMRDNRSELTYKYKTEPLDVAIPKARRVYOGEK RAYGALPLGFGAGSTMGCTSMTLTVAROLDLSIVQOQNNLRAEAQOHLQDL TVGICIKOLQRIILAVERYLKDQDLGLTWGSGSLICTTAPVNNASMSNLSLQDINNMN TMEWDEIEINNYLSLSLSEENQOGEDEBELLDKASLMNFNTINMYLTYKL FIMVIGVLQELRYFVAVLSIENNRQGSQPSLOTHTPIRGDRGPDJEEEGEGRD DRSIRLVNGSLALIMDLRSLCLFSYHRLRDLILIVRYIELLGRGWELKXMMNLL QXSGLEKNSAVNMLNATLAIVAEGRVLEVLQAAVRAIRHPRIRQGLERILL"
CDS	6221. .8785	8787. .9407
	6221. .8785	/codon_start=1
CDS	8787. .9407	/transl_table=11
	8787. .9407	/product="nef protein"
CDS	8787. .9407	/protein_id="AAK08490.1"
	8787. .9407	/db_xref="GI:12831143"
CDS	8787. .9407	/translation="MGKMSKSSVIGPVAVERMRRAEPADGCVGASVLDLEKHAIT SSNTAANNACAMLEAEDEEBEGVPVTPQVRLPMTYKAAVDSLHFLKREKGLGLIH SORRODILDLMTIYHOGYFPDMONITPGPVRPLTFEGMCYKLVPEPDVEANKGE NTSLSLHVSLHGHDDPDERVLEWRPDSRLAFHNVARELHPEYFKNC"
	8787. .9407	9076. .9709
CDS	9076. .9709	/transl_table=11
	9076. .9709	/note="3' long terminal repeat"
CDS	9076. .9709	9529. .9626
	9076. .9709	

```
/note="R: 3' copy"
polyA.signal      9602..9607
misc_feature      9710..11282
                  /note="3' cellular genomic sequence"
                  complement(112278..13138)
                  /note="from pUC vector"
CDS
  /codon_start=1
  /transl_table=11
  /product="beta-lactamase"
  /protein_id="AAK08491.1"
  /db_xref="GI:12831144"
  /translation="MSIOHFRVALIPFEAFCLPVFAHPETLVKVDADGAEVGY
  IELDLSKILSEFRPEPMSTFKYLICGAVLSIRDAEOLGRTHYSDNDIVE
  YSPYKELHLDGKAVRRELCSAITSMDTANLLITIGRKETAFILHMGDVTSL
  DRWPELNEATLPNDKDTMPVAAATLTKLLTGLLETLASRQDLIDMEDKVAQPL
  LRSALPWGFWFIADSGAGERSGRIIALGLPDKPSRLIVITVTSQATMDERNROIA
  EIGASLIKHM"
misc_feature      13647..14824
                  /note="5' cellular genomic sequence"
BASE COUNT      4733 a      3022 c      3607 g      3462 t
ORIGIN
Query Match      43.0%; Score 2676.2; DB 12; Length 14824;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      1291 CCATGAGGTAGAGAGAGATATCAGCACTCTGTGAGATGGGGGTGAGAAATGGGGCACC 1350
      |||||||
DB      6219 CAATGAGGTAGAGAGAGATATCAGCACTCTGTGAGATGGGGGTGAGAAATGGGGCACC 6278
OY      1351 TGTCTCTGGGATATTTGATGATCTGTAGTCTACAGAAATTTGGGGTCAACCGTCTATT 1410
      |||||||
DB      6279 TGTCTCTGGGATATTTGATGATCTGTAGTCTACAGAAATTTGGGGTCAACCGTCTATT 6338
OY      1411 ATGGGATCTGTGTGTGAGAGAGAACACCACTCTATTTTGTGTCATCAGATGCTAAAG 1470
      |||||||
DB      6339 ATGGGATCTGTGTGTGAGAGAGAACACCACTCTATTTTGTGTCATCAGATGCTAAAG 6398
OY      1471 CATATGATCAGAGGTACATATATGTTTGGGCGACACATGCGTGTACCCACAGACCCCA 1530
      |||||||
DB      6399 CATATGATCAGAGGTACATATATGTTTGGGCGACACATGCGTGTACCCACAGACCCCA 6458
OY      1531 ACCCAAGAGATGATTTGTTAAATGTGACAGAAATTTTAACTGTGAAAAATGACA 1590
      |||||||
DB      6459 ACCCAAGAGATGATTTGTTAAATGTGACAGAAATTTTAACTGTGAAAAATGACA 6518
OY      1591 TGTAGAGACAGATGATGATATATCAGTTATGAGGATCAAAAGCCTAAAGCCATGTG 1650
      |||||||
DB      6519 TGTAGAGACAGATGATGATATATCAGTTATGAGGATCAAAAGCCTAAAGCCATGTG 6578
OY      1651 TAAATTTAACCCTCTCTGTGTAGTTTAAAGTCACTGATTTGAGATGATCTAATA 1710
      |||||||
DB      6579 TAAATTTAACCCTCTCTGTGTAGTTTAAAGTCACTGATTTGAGATGATCTAATA 6638
OY      1711 CCAATAGTAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAAAGTCTCTTCA 1770
      |||||||
DB      6639 CCAATAGTAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAAAGTCTCTTCA 6698
OY      1771 ATATCAGCACAAGATAGATAGATGTCAGAAAAGATATGATTCATTTTATTAACCTG 1830
      |||||||
DB      6699 ATATCAGCACAAGATAGATAGATGTCAGAAAAGATATGATTCATTTTATTAACCTG 6758
OY      1831 ATATAGTACCAATAGATATATCCAGCTATAGGTGATATGTTGTAACCTCAGTCATTA 1890
      |||||||
DB      6759 ATATAGTACCAATAGATATATCCAGCTATAGGTGATATGTTGTAACCTCAGTCATTA 6818
OY      1891 CACAGGCTGTGCAAGGTATCTTTGAGCAATTCATCCATACATATATGTCGCCGGGTG 1950
      |||||||
DB      6819 CACAGGCTGTGCAAGGTATCTTTGAGCAATTCATCCATACATATATGTCGCCGGGTG 6878
OY      1951 GTTTTGCATTTCTAAATGTATATATAGATTCATATGATGAAACAGGACATGTACAATG 2010
      |||||||
DB      6879 GTTTTGCATTTCTAAATGTATATATAGATTCATATGATGAAACAGGACATGTACAATG 6938
```

```
OY      2011 TCAGCAGATACCAATGTATACATGGAATTCAGGCCAGTAGTATCACTCACTGCTTTAA 2070
      |||||||
DB      6939 TCAGCAGATACCAATGTATACATGGAATTCAGGCCAGTAGTATCACTCACTGCTTTAA 6998
OY      2071 ATGCAGTCTAGCAGAAAGATGATGATTAATAGATCTGCCAATTTTCACAGCAATGCTA 2130
      |||||||
DB      6999 ATGCAGTCTAGCAGAAAGATGATGATTAATAGATCTGCCAATTTTCACAGCAATGCTA 7058
OY      2131 AACCATATATAGTACAGCTGAACACATCTGTAGAATTAATTTGACAGAACCCCAACACA 2190
      |||||||
DB      7059 AACCATATATAGTACAGCTGAACACATCTGTAGAATTAATTTGACAGAACCCCAACACA 7118
OY      2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGAGACATTTGTTCAATAGGAA 2250
      |||||||
DB      7119 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGAGACATTTGTTCAATAGGAA 7178
OY      2251 AATATGAAATATGAGACACACATTTGTAATAGTAGAGCAAAATGSAATGCACTT 2310
      |||||||
DB      7179 AATATGAAATATGAGACACACATTTGTAATAGTAGAGCAAAATGSAATGCACTT 7238
OY      2311 TAAACAGATAGTACCAATTAAGAGAACAAATTTGAAATPATTAACCAATTAATCTTTA 2370
      |||||||
DB      7239 TAAACAGATAGTACCAATTAAGAGAACAAATTTGAAATPATTAACCAATTAATCTTTA 7298
OY      2371 AGCAATCCTCAGAGGGGGACCCAGAAATTTGTACGCCACAGTTTAAATTTGGAGGGGAT 2430
      |||||||
DB      7299 AGCAATCCTCAGAGGGGGACCCAGAAATTTGTACGCCACAGTTTAAATTTGGAGGGGAT 7358
OY      2431 TTTTCTACTGTAAATTCACACACACAGTAAATAGTACTGTGTTAAATAGTCTTGAGATA 2490
      |||||||
DB      7359 TTTTCTACTGTAAATTCACACACACAGTAAATAGTACTGTGTTAAATAGTCTTGAGATA 7418
OY      2491 CTGAAGGTCAAATTAACCTGAAAGAGTGAACACATCACCCTCCATGCAATTAAC 2550
      |||||||
DB      7419 CTGAAGGTCAAATTAACCTGAAAGAGTGAACACATCACCCTCCATGCAATTAAC 7478
OY      2551 AATTTATTAACATGCGGAGAGATGAGAAAGCAATGTATGCCCCCTCCATGATGTGAC 2610
      |||||||
DB      7479 AATTTATTAACATGCGGAGAGATGAGAAAGCAATGTATGCCCCCTCCATGATGTGAC 7538
OY      2611 AAATTAGATGTTCAATCAATATTTACTGGGCTGTATTAACAAGAGATGGTGTATTAACA 2670
      |||||||
DB      7539 AAATTAGATGTTCAATCAATATTTACTGGGCTGTATTAACAAGAGATGGTGTATTAACA 7598
OY      2671 ACAATGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGAGAAATG 2730
      |||||||
DB      7599 ACAATGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGAGAAATG 7658
OY      2731 AATTTATTAATTTATTAATAGTAAATTTGAACCTTTAGAGATGACACCCCAAGGCA 2790
      |||||||
DB      7659 AATTTATTAATTTATTAATAGTAAATTTGAACCTTTAGAGATGACACCCCAAGGCA 7718
OY      2791 AGAGAAGAGTGTGTCAGAGAGAAAAAAGAGAGTGGGAATAGAGACTTGTTCCTTGGGT 2850
      |||||||
DB      7719 AGAGAAGAGTGTGTCAGAGAGAAAAAAGAGAGTGGGAATAGAGACTTGTTCCTTGGGT 7778
OY      2851 TCTTGGAGACAGAGGAAGCACTATGGGCTCAGCTCAATGACGCTGACGGTACAGGCCA 2910
      |||||||
DB      7779 TCTTGGAGACAGAGGAAGCACTATGGGCTCAGCTCAATGACGCTGACGGTACAGGCCA 7838
OY      2911 GACATTTTGTCTGATTTAGTGCAGACGACGAACAAATTTGCTGAGGGCTATTGAGGGC 2970
      |||||||
DB      7839 GACATTTTGTCTGATTTAGTGCAGACGACGAACAAATTTGCTGAGGGCTATTGAGGGC 7898
OY      2971 AACAGCATCTGTGCACTCAGCTGAGGCGATCAACACACTCAGGCAAGAAATCTCGG 3030
      |||||||
DB      7899 AACAGCATCTGTGCACTCAGCTGAGGCGATCAACACACTCAGGCAAGAAATCTCGG 7958
OY      3031 CTGTGAAAGATATACCTAAAGGATCAACAGCTCTGGGGATTTGGGGTGTCTGTGAAAAAC 3090
      |||||||
DB      7959 CTGTGAAAGATATACCTAAAGGATCAACAGCTCTGGGGATTTGGGGTGTCTGTGAAAAAC 8018
```

QY	3091	TCATTTGGACACACCTCTGCTTGGATAGTCTAGTTGGAGATTAATTAATCTCTGGACACGA	3130
Db	8019	TCATTTGGACACACCTCTGCTTGGATAGTCTAGTTGGAGATTAATTAATCTCTGGACACGA	8078
QY	3151	TTTGGAAATACATGACCTGATGAGTGGAGACAGAAATTAACAAATTACACAGACTTAA	3210
Db	8079	TTTGGAAATACATGACCTGATGAGTGGAGACAGAAATTAACAAATTACACAGACTTAA	8138
QY	3211	TACACTCTCTTAATTGSAACAATCGCAAAACACGACAAAGAAATGACACAGATTTATTGG	3270
Db	8139	TACACTCTCTTAATTGSAACAATCGCAAAACACGACAAAGAAATGACACAGATTTATTGG	8198
QY	3271	AATTAGATAAATGGGCAAGTTTGTTGGAATTTGGTTTAACATTAACAATTTGGCTGTGATA	3330
Db	8159	AATTAGATAAATGGGCAAGTTTGTTGGAATTTGGTTTAACATTAACAATTTGGCTGTGATA	8258
QY	3331	TAAATTTATCATTAATGATAGTAGAGGCTTGTAGTGTTAAGAAATAGTTTTCGTGTAC	3390
Db	8259	TAAATTTATCATTAATGATAGTAGAGGCTTGTAGTGTTAAGAAATAGTTTTCGTGTAC	8318
QY	3391	TTTCATATGTAATTAAGATTTAGGACAGGATATTAACCATATTCGTTTAGACCACTCC	3450
Db	8319	TTTCATATGTAATTAAGATTTAGGACAGGATATTAACCATATTCGTTTAGACCACTCC	8378
QY	3451	CAATTCGAGGAGGACCCGACAGGCCCGGAGGAATAGAAAGAAAGTGGAGAGAGACA	3510
Db	8379	CAATTCGAGGAGGACCCGACAGGCCCGGAGGAATAGAAAGAAAGTGGAGAGAGACA	8438
QY	3511	GAGACAGATTCATTTCGATTAGTGAACGATCTCTTAGCACTTATCTGGAGCATCTCGGA	3570
Db	8439	GAGACAGATTCATTTCGATTAGTGAACGATCTCTTAGCACTTATCTGGAGCATCTCGGA	8498
QY	3571	GCCATGCGCTCTTCAGCTACCAACGCTTGAAGAGACTTACTCTTGATTGTACAGAGATTG	3630
Db	8499	GCCATGCGCTCTTCAGCTACCAACGCTTGAAGAGACTTACTCTTGATTGTACAGAGATTG	8558
QY	3631	TGGAATCTCTGGAGACGCGAGGGGGTGGAAACCCCTCAATATTGGTGAATCTCTACAGT	3690
Db	8559	TGGAATCTCTGGAGACGCGAGGGGGTGGAAACCCCTCAATATTGGTGAATCTCTACAGT	8618
QY	3691	AATGAGATCAGGAATAAGCAATATAGTGCCTGTTAACTTCTCAATGCCACAGCCATAGCAG	3750
Db	8619	AATGAGATCAGGAATAAGCAATATAGTGCCTGTTAACTTCTCAATGCCACAGCCATAGCAG	8678
QY	3751	TAGCTGAGGGGACAGATAGGGTTATAGCAATTTACAAGCAGCTTATAGACTATTGGCC	3810
Db	8679	TAGCTGAGGGGACAGATAGGGTTATAGCAATTTACAAGCAGCTTATAGACTATTGGCC	8738
QY	3811	ACATATCCATAGAAATAGACAGGCGCTTGAAGAGATTTTCTTAAAGATGGTGGCAG	3870
Db	8739	ACATATCCATAGAAATAGACAGGCGCTTGAAGAGATTTTCTTAAAGATGGTGGCAG	8798
QY	3871	TGTCCTAAAAAGTAGTGTGATTTGGATGGCCCTCTTAAAGGAAAGATGACAGACTGAG	3930
Db	8799	TGTCCTAAAAAGTAGTGTGATTTGGATGGCCCTCTTAAAGGAAAGATGACAGACTGAG	8858
QY	3931	CCAGCAGCAGATGGGGTGGGACGAGTATCTCGAATCTAGACCTAGA	3971
Db	8859	CCAGCAGCAGATGGGGTGGGACGAGTATCTCGAATCTAGACCTAGA	8899
RESULT 7			
LOCUS	HIV1U26942	9000 bp	DNA
DEFINITION	Human immunodeficiency virus clone pNL4-3, subclone 4.20.	linear	VRL 15-NOV-1995
VERSION	U26942		
KEYWORDS	U26942.1	GI:902798	
SOURCE	Human immunodeficiency virus 1 (HIV-1)		
ORGANISM	Human immunodeficiency virus 1 (HIV-1)		
REFERENCE	1 (bases 1 to 9000)		

AUTHORS	Salminen,M.O., Koch,C., Sanders-Buell,E., Ehrenberg,P.K., Michael,N.L., Carr,J.K., Burke,D.S. and McCutchan,F.E.
TITLE	Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction
JOURNAL	Virology 213 (1), 80-86 (1995)
MEDLINE	960366482
REFERENCE	7483282
AUTHORS	2 (bases 1 to 9000)
TITLE	Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R., Rabson,A. and Martin,M.A.
JOURNAL	Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone
MEDLINE	J. Virol. 59 (2), 284-291 (1986)
REFERENCE	86281827
AUTHORS	3016298
TITLE	3 (bases 1 to 9000)
JOURNAL	Salminen,M.S.
REFERENCE	Direct Submission
AUTHORS	Submitted (11-MAY-1995) Mka S. Salminen, Program area 1, Henry M. Jackson Foundation, 1600 East Gude Drive, Rockville, MD 20850, USA
TITLE	Location/Qualifiers
FEATURES	1..9000
source	/organism="Human immunodeficiency virus 1"
	/mol_type="genomic DNA"
	/strain="NL4-3"
	/db_xref="taxon:11676"
	/note="Long PCR amplified, re-cloned 9kb fragment from infectious HIV-1 clone pNL4-3, subclone 4.20. Insert sequence only, vector is pNOTA. Cloned into Invitrogen vector pNOTA from extended PCR amplification product using original pNL4-3 molecular clone as template. Only 17 mutations compared to original published sequence (Genbank Accession Number M19921) were observed."
LTR	<1..13
	/note="5' LTR"
gene	168..1670
	/gene="gag"
CDS	168..1670
	/gene="gag"
	/note="matrix (p17), capsid (p24) and core (p7, p6) proteins"
	/codon_start=1
	/product="Gag polyprotein precursor"
	/protein_id="AAB60571.1"
	/db_xref="GI:902799"
	/translation="MGARASVLSGGELDKMKETRLRPGKROYKRLKHIWASLELRF AVNPGILTESEGGKILGOLPDISLSEELRSYNTIAYLCVHORIDVDRKALD KIEEONKSKKKAQQAADPTGNNSSQVQNPYQNDQGVYHQAISRTINAAKYVE EKAPSPVITPMFSALSBGATPQDNLNTLVGGHQAQMFKETINBEAEWRLLPV HAGIAGQGRPEPRGSDIAGTSTLQDQIGMWHNPETPVGEITYKRWIITGLKRIIVM YSPSTLIDIRQGRKEPRRDYVDRFYKTLRAEQASOEKVMNMFETLLVQANPCKTIL KALPGSALTEEMWTACQGVGPGHKAVLAEAMSQVYNPATIMIQGKFNQKRTYVC FNGKKECHIAKNCRAPKCKGCMKCGKRGHMKQDQANFLKIKPMSHKGGRNLFQ SRPEPTAPPEESRFGEETTPPQGRQRPIDKELXPLASLRLSGDSSQ"
conflict	1176
	/gene="gag"
	/citation=[2]
	/replace="t"
gene	1463..4474
	/gene="pol"
	<1463..4474
CDS	/gene="pol"
	/note="protease, RT, RNase H, Integrase; Translated by -1 frameshift at TT codon at end of gag p7"
	/codon_start=1
	/product="POL polyprotein precursor"
	/protein_id="AAB60572.1"
	/db_xref="GI:902800"
	/translation="FFREDLAFPGKAREFSSBOTRANSPTRRELQVWGNDNLSSEA GARGQVSPSPQITLMQRPVLTIKIGGGLKALLDGTGADTVLSEMNIPGKMKPKM IGGGGIKTRYQDOLIEICGKAITGVVGTPVNIIGRNLTQIGCTINPISPI

```
ETVPYKLGKMGDPKVKOMPLTEERIKALVEICTEMEREKISKIGPENPNYNTFVAI
KKKDKTKRKLIVDFRELNRKRTODEFWEVOLGIPHPAGLKKOKSVTLVDDGDAFESVLD
KDFRKYTAFTTIPSTINNENPGTROYVNVLPDGMKSSPAIFCCSMPTILPFPKRONPDVY
IYQYMDDIYVGSDELTEIGORHKIEBLRHLRMGTTTDDKHKQKPEPLLMGCEYELHPD
KWTYQPIVLPEKDSMTVNDIOKLVLGKLMWASQITAGIVROLCKLIRGTKALTEYVPL
TEAELELAENNEELKEPHGVYDPSKDLAELOKOGQMTYQIYDEPRLNLTGK
YARMKGATNDVKOLTEAVOKIATESIYVWGTPFKLPIOKETWEAMTWEQWATVI
PEMEFVNTPPVLKMYOLEKEPIIGAEFEYVYGANRETKLGKAGYVTDGROKVVPL
TDTMOKTELOAIHIALODSGLEVINIYDISOYALGIIQAQDPKSESEIYSOTIEBLIK
KEKYLLAWPAHAKGTGNEOYDKLVASGIRKVLFDIGDKAOEHEKTHSMRWASD
ENLPPVNAKEIVASDCOLKGEAMHGOVDSPIGWLDCTHLEGVILVAVHVASGY
IEAEVIPAETGOETAYFLKLAGRWPKVTHDNGSNFTSTVKAACWAGIKQFEGI
PYNPOSQVIESMKNELKIIIGOVDOAEHLKTAVQMAVFTHNEFRKGIGISYAGER
IVDIATDIOTKELOKLTITKINFEVYRDSRDPVWKGPAKLWKGEGAVYIQDNDSI
KVVPRRKAKIIRIDYCKOMAGDCCVARSRODED"
1591
/gene="gag"
/citation=[2]
/replace="a"
1799
/gene="pol"
/citation=[2]
/replace="g"
2459
/gene="pol"
/citation=[2]
/replace="a"
2986
/gene="pol"
/citation=[2]
/replace="a"
3575
/gene="pol"
/citation=[2]
/replace="g"
3576
/gene="pol"
/citation=[2]
/replace="g"
3577
/gene="pol"
/citation=[2]
/replace="g"
3832
/gene="pol"
/citation=[2]
/replace="g"
4009
/gene="pol"
/citation=[2]
/replace="a"
4119
/gene="vif"
4419
/gene="vif"
/codon_start=1
/product="vif protein"
/protein_id="AAB60573.1"
/db_xref="GI:902801"
/gene="vpr"
4937
/gene="vpr"
4937
/gene="vpr"
/codon_start=1
/product="vpr protein"
/protein_id="AAB60574.1"
/db_xref="GI:902802"
/translation="MEQAPEDQGPQREPYNEMTLELEELKSEAVRHPRJLWNLGR
HIYETGDTMAGVEALIRIQLPFIHRIICGRHSRISGVTRRRARNCASRS"
```

```
conflict 4981
/gene="vpr"
/citation=[2]
/replace="t"
5067
/gene="vpr"
/citation=[2]
/replace="a"
5139
/gene="vpr"
/citation=[2]
/replace="c"
5208
/gene="tat"
join(5208..5422,7747..7792)
/gene="tat"
/function="transactivator"
/codon_start=1
/product="TAT protein"
/protein_id="AAB60575.1"
/db_xref="GI:902803"
/translation="MEVPDPLPEPKHPGSQPKTACTNCTCKKCCFHQCVCFTKALG
ISYGRKRORRRRAHONSOTHOASLSKOPTSOSRDPGPKE"
5347
/gene="rev"
join(5347..5422,7747..8021)
/gene="rev"
/function="regulator of viral expression"
/codon_start=1
/product="REV protein"
/protein_id="AAB60576.1"
/db_xref="GI:902804"
/translation="MAGRSGSDSELIPTVPLIKLLYOSNPPNPECTRQARRRRR
WRERQRIHSISERLSTYIGRSEVPVLQPLERLITLDCNEDCGTSGTGVSPPQI
LVSEPTVLESCTKE"
5439
/gene="vpu"
5439
/gene="vpu"
/codon_start=1
/product="vpu protein"
/protein_id="AAB60577.1"
/db_xref="GI:902805"
/translation="MQPIIVAVLVAIIIAIYVWSIIVIEYKILRQKIDRLIDR"
42.8% Score 2665; DB 14; Length 9000;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2671; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1291 CCATGAGAGTAAGAGAGATATCAGCACTGTGAGATGGGGGGAATGGGCACCA 1350
DB 5597 CAATGAGAGTAAGAGAGATATCAGCACTGTGAGATGGGGGGAATGGGCACCA 5656
QY 1351 TGCCTCTGGGATATGATGATCTGTAGTGTACAGAAAATGTGGGTCAACCGTATT 1410
DB 5657 TGCCTCTGGGATATGATGATCTGTAGTGTACAGAAAATGTGGGTCAACCGTATT 5716
QY 1411 ATGGGCTACCTGTGTGAAGAGACACCACTCTATTTTGTGCATGATGCTTAAG 1470
DB 5717 ATGGGCTACCTGTGTGAAGAGACACCACTCTATTTTGTGCATGATGCTTAAG 5776
QY 1471 CATATGATACAGAGTACATATATTTGGGCGACACATGCCGTGTACCCAGACGCCA 1530
DB 5777 CATATGATACAGAGTACATATATTTGGGCGACACATGCCGTGTACCCAGACGCCA 5836
QY 1531 ACCCACAAGTAGTATTTGTAATGTGACAGAAAATTTTAACATGTGAAAATGACA 1590
DB 5837 ACCCACAAGTAGTATTTGTAATGTGACAGAAAATTTTAACATGTGAAAATGACA 5896
QY 1591 TGTGTAACAGATGATGATATATATGATTTATGGGATCAAGCCTTAAGCCTATGTG 1650
DB 5897 TGTGTAACAGATGATGATATATATGATTTATGGGATCAAGCCTTAAGCCTATGTG 5956
```

1651 TAAATTAACCCACTGTGTGTAGTTAAAGTGCACATGATTTGAAGATGATTAATA 1710  
| | | | |  
Db TAAATTAACCCACTGTGTGTAGTTAAAGTGCACATGATTTGAAGATGATTAATA 6016  
5957  
1711 CCAATAGTAGTGGCGGGAATGATTAATGAGAAAGAGAGATTAATAAACTGCTTTCA 1770  
| | | | |  
Db CCAATAGTAGTGGCGGGAATGATTAATGAGAAAGAGAGATTAATAAACTGCTTTCA 6076  
6017  
1771 ATATCAGACAAGCATAGAGATTAAGTGCAGAAAGATATGCACTTTCTTTTAAACTTG 1830  
| | | | |  
Db ATATCAGACAAGCATAGAGATTAAGTGCAGAAAGATATGCACTTTCTTTTAAACTTG 6136  
6077  
1831 ATATGATGCATATAGATTAATACAGCTATAGTGTATTAAGTTGAACCTCAGTCATTA 1890  
| | | | |  
Db ATATGATGCATATAGATTAATACAGCTATAGTGTATTAAGTTGAACCTCAGTCATTA 6196  
6137  
1891 CACAGGCCGTGCCAAGGATTCCTTGAGCCAATCCCATATATATGTGCCCCGGCTG 1950  
| | | | |  
Db CACAGGCCGTGCCAAGGATTCCTTGAGCCAATCCCATATATATGTGCCCCGGCTG 6256  
6197  
1951 GTTTTGCATCTTAAATATGATTAATTAAGACGTTCAATGAAACAGACCATATCAATG 2010  
| | | | |  
Db GTTTTGCATCTTAAATATGATTAATTAAGACGTTCAATGAAACAGACCATATCAATG 6316  
6257  
2011 TCAGCAGCTACATATGTACATGGAATCAGGCCAGTATCAACTCACTGCTTTAA 2070  
| | | | |  
Db TCAGCAGCTACATATGTACATGGAATCAGGCCAGTATCAACTCACTGCTTTAA 6376  
6317  
2071 ATGGCAGCTCTAGCAGAAAGATGATTAATTAAGTGCACCAATTCACAGCATATGCTA 2130  
| | | | |  
Db ATGGCAGCTCTAGCAGAAAGATGATTAATTAAGTGCACCAATTCACAGCATATGCTA 6436  
6377  
2131 AAACCATATATAGTACAGCTGAACACATCTGTAGAATTAATTTGACAAGCCCAACA 2190  
| | | | |  
Db AAACCATATATAGTACAGCTGAACACATCTGTAGAATTAATTTGACAAGCCCAACA 6496  
6437  
2191 ATACAGAAAAAGTTCGGTATCCAGAGGGGCCAGGAGAGCATTTGTTCAATAGGAA 2250  
| | | | |  
Db ATACAGAAAAAGTTCGGTATCCAGAGGGGCCAGGAGAGCATTTGTTCAATAGGAA 6556  
6497  
2251 AAATGGAATATGAGACAGACACATTTAATCATTTAGTAGAGCAAAATGGAATGCCATT 2310  
| | | | |  
Db AAATGGAATATGAGACAGACACATTTAATCATTTAGTAGAGCAAAATGGAATGCCATT 6616  
6557  
2311 TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGAAATTAATTAATTAATCTTTA 2370  
| | | | |  
Db TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGAAATTAATTAATTAATCTTTA 6676  
6617  
2371 AGCATCCTCAGAGGGGAGACCAGAAATTTGAACCCACAGTTTAATTTGTGAGGGGAAT 2430  
| | | | |  
Db AGCATCCTCAGAGGGGAGACCAGAAATTTGAACCCACAGTTTAATTTGTGAGGGGAAT 6736  
6677  
2431 TTTTCTACTGTAATTCACACACACATGTTAATAGTACTGTTTAAATAGTACTTGGAGTA 2490  
| | | | |  
Db TTTTCTACTGTAATTCACACACACATGTTAATAGTACTGTTTAAATAGTACTTGGAGTA 6796  
6737  
2491 CTGAAGGCTCAATTAACACTGAAGAGTGAACACATCACACTCCCATGAGAAATTAAC 2550  
| | | | |  
Db CTGAAGGCTCAATTAACACTGAAGAGTGAACACATCACACTCCCATGAGAAATTAAC 6856  
6797  
2551 AATTTATTAACATGTGGCAGAGATGAGAAAAAGCATATATGCCCCCTCCATCAGTGCAC 2610  
| | | | |  
Db AATTTATTAACATGTGGCAGAGATGAGAAAAAGCATATATGCCCCCTCCATCAGTGCAC 6916  
6857  
2611 AATATAGATGTTCAATTAATTTACTGGGCTGCTATTAAACAAGAGATGCTGTATTAACA 2670  
| | | | |  
Db AATATAGATGTTCAATTAATTTACTGGGCTGCTATTAAACAAGAGATGCTGTATTAACA 6976  
6917  
2671 ACAATGGCTCGAGATCTTCAGACCTGAGGAGGCGATATGAGGACATTTGAGAAATG 2730  
| | | | |  
Db ACAATGGCTCGAGATCTTCAGACCTGAGGAGGCGATATGAGGACATTTGAGAAATG 7036  
6977  
2731 AATTAATTAATTAATTAAGTAAATAAATTTGAACCATTAAGAGTAGCACCACCAAGCAA 2790  
| | | | |  
Db AATTAATTAATTAATTAAGTAAATAAATTTGAACCATTAAGAGTAGCACCACCAAGCAA 2790

7037 AATTAATTAATTAATTAAGTAAATAAATTTGAACCATTAAGAGTAGCACCACCAAGCAA 7096  
| | | | |  
Db AATTAATTAATTAATTAAGTAAATAAATTTGAACCATTAAGAGTAGCACCACCAAGCAA 7096  
7037  
2791 AGAGAAAGTGTGTCAGAGAGAAAAAAGACAGTGGGGAATAGAGCTTTGTTCTTGCGT 2850  
| | | | |  
Db AGAGAAAGTGTGTCAGAGAGAAAAAAGACAGTGGGGAATAGAGCTTTGTTCTTGCGT 7156  
7097  
2851 TCTTGGGAGCAGCAGGAAGCACTATGAGCTCAGTCAATAGACGCTGACGGTACAGGCCA 2910  
| | | | |  
Db TCTTGGGAGCAGCAGGAAGCACTATGAGCTCAGTCAATAGACGCTGACGGTACAGGCCA 7216  
7157  
2911 GACATTAATTTGTCGATATAGTGCACAGCAGCAAAATTTGCTGAGGCTATTTAGGCGC 2970  
| | | | |  
Db GACATTAATTTGTCGATATAGTGCACAGCAGCAAAATTTGCTGAGGCTATTTAGGCGC 7276  
7217  
2971 AACAGCATCTGTGCACTACAGTCTGGGCACTCAAAACACTCCAGCAGAAATCCGCG 3030  
| | | | |  
Db AACAGCATCTGTGCACTACAGTCTGGGCACTCAAAACACTCCAGCAGAAATCCGCG 7336  
7277  
3031 CTGTGAAAGATTAACCTTAAGAGATCAACAGCTCTGCGGATTTGGGTTGCTGTGAAAC 3090  
| | | | |  
Db CTGTGAAAGATTAACCTTAAGAGATCAACAGCTCTGCGGATTTGGGTTGCTGTGAAAC 7396  
7337  
3091 TCATTTGCACACAGCTGCTGCTTGGAAATGCTTGTGAGTAAATTAATCTGTGACAGCA 3150  
| | | | |  
Db TCATTTGCACACAGCTGCTGCTTGGAAATGCTTGTGAGTAAATTAATCTGTGACAGCA 7456  
7397  
3151 TTTGGAATTAACATGACCTGATGAGTGGGAGAGGAATTAACAAATTAACAGCTTTAA 3210  
| | | | |  
Db TTTGGAATTAACATGACCTGATGAGTGGGAGAGGAATTAACAAATTAACAGCTTTAA 7516  
7457  
3211 TACACTCCTTAATTAAGAAATGCAAAACAGCAAGAAAGAAATGAACAAATTAATTTGG 3270  
| | | | |  
Db TACACTCCTTAATTAAGAAATGCAAAACAGCAAGAAAGAAATGAACAAATTAATTTGG 7576  
7517  
3271 AATTAATTAATTAAGGCAAGTTTGGAAATGCTTAATTAACAAATTTGCTGTGAGTAA 3330  
| | | | |  
Db AATTAATTAATTAAGGCAAGTTTGGAAATGCTTAATTAACAAATTTGCTGTGAGTAA 7636  
7577  
3331 TAAATTAATTAATTAAGTATAGTAGAGGCTTGTGATTAAGAAATGTTTCTGTAC 3390  
| | | | |  
Db TAAATTAATTAATTAAGTATAGTAGAGGCTTGTGATTAAGAAATGTTTCTGTAC 7696  
7637  
3391 TTTCTATATGTAATAGCTTAGAGCAGGGAATTCACATTAATGCTTACAGCCACCTCC 3450  
| | | | |  
Db TTTCTATATGTAATAGCTTAGAGCAGGGAATTCACATTAATGCTTACAGCCACCTCC 7756  
7697  
3451 CAATCCGAGGGGAGCCGACAGGCCGGAAGGAATGAAGAAAGAGTGGAGAGAGCA 3510  
| | | | |  
Db CAATCCGAGGGGAGCCGACAGGCCGGAAGGAATGAAGAAAGAGTGGAGAGAGCA 7816  
7757  
3511 GAGACAGATCCATTCGATTAAGTGAACGATCTTAAGCACTATCTGGAGCATCTGCGGA 3570  
| | | | |  
Db GAGACAGATCCATTCGATTAAGTGAACGATCTTAAGCACTATCTGGAGCATCTGCGGA 7876  
7817  
3571 GCTGTGCTCTTACGCTTACACCGCTTGAGAGACTTACCTGATTTGAAGAGATG 3630  
| | | | |  
Db GCTGTGCTCTTACGCTTACACCGCTTGAGAGACTTACCTGATTTGAAGAGATG 7936  
7877  
3631 TGGAACTTCTGGGACGAGGGGGTGGGAAGCCCAATATATTTGGTGAATCTCTACAGT 3690  
| | | | |  
Db TGGAACTTCTGGGACGAGGGGGTGGGAAGCCCAATATATTTGGTGAATCTCTACAGT 7996  
7937  
3691 ATTGGAGTCAGAGCTAAGAAATAGTGTATTAATCTGCTCAATGCCCACAGCATATAGG 3750  
| | | | |  
Db ATTGGAGTCAGAGCTAAGAAATAGTGTATTAATCTGCTCAATGCCCACAGCATATAGG 8056  
7997  
3751 TAGCTGAGGGGAGAGATAGGCTTATAGAGTATTAACAAGCAGCTTATAGAGCTATTTGCC 3810  
| | | | |  
Db TAGCTGAGGGGAGAGATAGGCTTATAGAGTATTAACAAGCAGCTTATAGAGCTATTTGCC 8116  
8057  
3811 ACATTACTAGAGAAATTAAGACAGGCTTGAAGAGATTTGCTATTAAGATGGTGGCAG 3870  
| | | | |  
Db ACATTACTAGAGAAATTAAGACAGGCTTGAAGAGATTTGCTATTAAGATGGTGGCAG 3870



[illegible]

CDS		/gene="env"
		6260..8830
		/gene="env"
		/codon.start=1
		/product="Env"
		/protein_id="BA013003.1"
		/db_xref="GI:1398972"
		/translation="MRVEKRYOHLMBWMKMTLGLMLICSTKELMTVTVYGVV WKATPTLPCASDAKADYDEVHNVWTHACVPDPNFOEVVLVAVTNEPMKMDAYE OMHEDITSLMDOSLPCVKLPLCVSLKCTKDKNDKNTSSSGMIENKEIKCSN ISTSKRGKVKEXEAFPEKLDIIPIDNFTSYLTCSITVITQACPVSEPIPIHC APAGFAILKCNKNTFNKTPCTVSTVQCTHGIRPVSTIOLINGSLAEBAVIRSN FTDNAKTIVOLTVSEINCTRPNNNRKRIORGGRAPVITGKIGNMKAQCNIS RAKMNATLKQIASKLREDFGNNTIIFKOSGGDEPEIVTSHFNGGEFYCNSTOLF STWNSTWTEGSSNNEGSDITLPCRIKOFIMMOVKRAMYAPISGQIRCSNIT GLLITRDGNSNNESEIFRPGGDMRDMNSLEYKYKVVLEIPLGAPRKAKRRVOR EKRAVIGALPFIPLGASGTMGCTSMITLVQAROLLISGIVQOONILRAIEAOHL OLTVWGIKOLQARLAVERYLKDOOLINGCSGLICTTAIVPNNAWSMKSLEQIWN NMTMEMDREINNTSLIHLIESONOQKNEDELLEDKMSLMMWFNTIMLWYI KLFIMIVGVLGRLIVEAFVSVNVRGOGYSPISFOYHLPPTPRGDPDEGLEEGER DRDRSIRLVNGLALIMDDLRITCLFXYHRLRDLILVTRIVELIGRGEMALKYMN LQYVSEGLKNSAVSLINATAINAECDTAVIEVQDAVIRIIRIRINOGLERILL "
repeat_region	9119..9754	
BASE COUNT	3432 a 1784 c 2372 g 2166 t	
ORIGIN		
Query Match	41.9%; Score 2610.6; DB 14; Length 9754;	
Best Local Similarity	98.5%; Pred. No. 0;	
Matches 2647; Conservative	0; Mismatches 34; Indels 6; Gaps 1;	
QY	1291 CCATGAGAGTGAAGAGAGATATCAGACCTTGTGAGATGGGGGTGGAATGGGGCACCA	1350
DB	6258 CAATGAGAGTGAAGAGAGAAATATCAGACCTTGTGAGATGGGGGTGGAATGGGGCACCA	6317
QY	1351 TGCCTCTTGGGATTTATGATGCTGTGTGTACAGAAAATTTGGGTACCGTCTAAT	1410
DB	6318 TGCCTCTTGGGATTTATGATGCTGTGTGTACAGAAAATTTGGGTACCGTCTAAT	6377
QY	1411 ATGGGCTACCTGTGTGAGAGAACCAACCACTATTTTGGCATTCAGATCTTAAG	1470
DB	6378 ATGGGCTACCTGTGTGAGAGAACCAACCACTATTTTGGCATTCAGATCTTAAG	6437
QY	1471 CATATGATACAGAGGTACATTAATTTGGGCGACACATGCTGTGTACCCACAGACCCA	1530
DB	6438 CATATGATACAGAGGTACATTAATTTGGGCGACACATGCTGTGTACCCACAGACCCA	6497
QY	1531 ACCCACAAGATGATTTGTGTAATGTGACAGAAAATTTTACATGTGGAATAATGACA	1590
DB	6498 ACCCACAAGATGATTTGTGTAATGTGACAGAAAATTTTACATGTGGAATAATGACA	6557
QY	1591 TGGTAGAAGATGATGAGAGATTAATCAAGTTATGGGATCAAGGCTAAAGCCATGTG	1650
DB	6558 TGGTAGAAGATGATGAGAGATTAATCAAGTTATGGGATCAAGGCTAAAGCCATGTG	6617
QY	1651 TAAATTTAACCCCACTGTGTAGTTAAAGTGCACATGATTTGGAAGATGATCTATA	1710
DB	6618 TAAATTTAACCCCACTGTGTAGTTAAAGTGCACATGATTTGGAAGATGATCTATA	6677
QY	1711 CCAATAGTATGAGGGAGATGATTAATGAGAAAGAGAGATTAATAAAGCTCTTTTCA	1770
DB	6678 CCAATAGTATGAGGGAGATGATTAATGAGAAAGAGAGATTAATAAAGCTCTTTTCA	6737
QY	1771 ATATCAGCAGAGATTAAGATTAAGTGCAGAAAGATTAATGCACTCTTTTAACTTG	1830
DB	6738 ATATCAGCAGAGATTAAGATTAAGTGCAGAAAGATTAATGCACTCTTTTAACTTG	6797
QY	1831 ATATAGTACCAATGATNA-----TACCACTATAGCTGATTAAGTTGTAACACCTCAG	1884
DB	6798 ATATAGTACCAATGATTAAGTACCTACAGCTATACGTAAGTAAGTTGTAACACCTCAG	6857
QY	1885 TCATTTACAGAGGCTGTCCAAAGGTATCTTTTGAGCCAAATTCATATTAATTTGTCCC	1944

DB	6858 TCATTTACAGAGGCTGTCCAAAGGTATCTTTTGAGCCAAATTCATATTAATTTGTCCC	6917
QY	1945 CGGCTGTTTGGCATCTTAATAATGTAATATAGACGTTCAATGGAACGAGACCATGTA	2004
DB	6918 CGGCTGTTTGGCATCTTAATAATGTAATATAGACGTTCAATGGAACGAGACCATGTA	6977
QY	2005 CAAATGTACAGCAGTACAAATGTAACATGGAATCAGGCCAGTATGATCACTCACTGC	2064
DB	6978 CAAATGTACAGCAGTACAAATGTAACATGGAATGAGCCAGTATGATCACTCACTGC	7037
QY	2065 TGTTAATGGCAGCTGTGAGAGAGAAATATGTAATTAATCTGCCAATTTTACAGACA	2124
DB	7038 TGTTAATGGCAGCTGTGAGAGAGAAATATGTAATTAATCTGTCAATTTTACAGACA	7097
QY	2125 ATGCTAAACCATATATGTCACGTGAACCATCTGTAGAATTAATTTAGCAAGACCA	2184
DB	7098 ATGCTAAACCATATATGTCACGTGAACCATCTGTAGAATTAATTTAGCAAGACCA	7157
QY	2185 ACAACAATACAGAGAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTACAA	2244
DB	7158 ACAACAATACAGAGAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTACAA	7217
QY	2245 TAGGAAAATATGGAATATATGAGACACACATTTGTAATAGTACGCAAAATGCAATG	2304
DB	7218 TAGGAAAATATGGAATATATGAGACACACATTTGTAATAGTACGCAAAATGCAATG	7277
QY	2305 CCACCTTAATAACAGATAGTACGCAAAATTAAGAGAACTTTGGAATATTAATAACATTA	2364
DB	7278 CCACCTTAATAACAGATAGTACGCAAAATTAAGAGAACTTTGGAATATTAATAACATTA	7337
QY	2365 TCTTTAAGCAATCCTCAGAGAGGGGACCAAGAAATTTGTAAGCAGCATTTTAATTTGGAG	2424
DB	7338 TCTTTAAGCAATCCTCAGAGAGGGGACCAAGAAATTTGTAAGCAGCATTTTAATTTGGAG	7397
QY	2425 GGGAAATTTTCTACTGTAATTTCAACACAACTGTTTAATAGTACTGTTTAATAGTACTT	2484
DB	7398 GGGAAATTTTCTACTGTAATTTCAACACAACTGTTTAATAGTACTGTTTAATAGTACTT	7457
QY	2485 GGAATGACTGAAGGCTCAATTAACCTGAGAGAAAGTGAACAACTCCATCCATCGAA	2544
DB	7458 GGAATGACTGAAGGCTCAATTAACCTGAGAGAAAGTGAACAACTCCATCCATCGAA	7517
QY	2545 TAAACAATTTATTAACATGTGCGAGAGATGAGAAAAGCAATGTATGCCCTCCATCA	2604
DB	7518 TAAACAATTTATTAACATGTGCGAGAGATGAGAAAAGCAATGTATGCCCTCCATCA	7577
QY	2605 GTGACAAATTTAGATGTTTCATCAATATTACTGGGCTGCTATTTAACAAAGATGTGTGTA	2664
DB	7578 GTGACAAATTTAGATGTTTCATCAATATTACTGGGCTGCTATTTAACAAAGATGTGTGTA	7637
QY	2665 ATTAACAACATGGGTCCGAGATCTTCAGACCTGAGAGAGGCGATATGAGGAGACAATTTGA	2724
DB	7638 ATTAACAACATGGGTCCGAGATCTTCAGACCTGAGAGAGGCGATATGAGGAGACAATTTGA	7697
QY	2725 GAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2784
DB	7698 GAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	7757
QY	2785 AGGCAAGAGAGAGATGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2844
DB	7758 AGGCAAGAGAGAGATGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	7817
QY	2845 TTGGGTTCTTGGGAG	2904
DB	7818 TTGGGTTCTTGGGAG	7877
QY	2905 AGGCGACACATTAATTTCTGATATATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2964
DB	7878 AGGCGACACATTAATTTCTGATATATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	7937
QY	2965 AGGCGACACAGCATCTTTGCAACTCAGACTGTGGGCTGCAAAAGAGCTCCAGGCAAGAA	3024
DB	7938 AGGCGACACAGCATCTTTGCAACTCAGACTGTGGGCTGCAAAAGAGCTCCAGGCAAGAA	7997



gene  
CDS  
5041..5619  
/gene="vif"  
5041..5619  
/gene="vif"  
/codon\_start=1  
/product="vif protein"  
/protein\_id="AAC28447.1"  
/db\_xref="GI:3378124"  
/translation="MENRMOVMYMOVDRMRINWKRLVYKHMTSKAKDMFYRHH  
ESTNPKISSEVHPILGDAKLVIYTWGLHGERMHLGQVSIEMRRKRRSTQVDPDL  
ADQLIHAFDCESESARNTILIRIVSPCEYAGHNKVGSLQYLALAIKPKQIK  
PLPSVRKLTEDRMNKPDKTGRHSHTMNH"

gene  
CDS  
5559..5729  
/gene="vpr"  
5559..5729  
/codon\_start=1  
/product="truncated vpr protein"  
/protein\_id="AAC28448.1"  
/db\_xref="GI:3378125"  
/translation="MEQAPEDGQRPREPYMENTLELLELSEAVRHFPRIIMHLSM  
KLTCILGDEMKO"  
5685-5686  
/gene="vpr"  
/note="site of difference between this sequence and that  
of clone NL4-3 (GenBank Accession Number M19921)"  
/replace="GGACACATA"  
5820..8404  
/gene="tat"  
join(5820..6034,8359..8404)  
/gene="tat"  
/codon\_start=1  
/product="tat protein"  
/protein\_id="AAC28449.1"  
/db\_xref="GI:3378126"  
/translation="MEPVDPRLPEPMKHGSPKTYACTNCCYKCKCFHCOCVFMTKALG  
ISYGRKRROBRRAHONSQTHQASLKSQPTSQSGRDPGPKF"  
5959..8633  
/gene="rev"  
join(5959..6034,8359..8633)  
/gene="rev"  
/codon\_start=1  
/product="rev protein"  
/protein\_id="AAC28450.1"  
/db\_xref="GI:3378127"  
/translation="MAGRSGDSDELLRTVPLIKILYQSNPPPEPTEGROARRRRR  
WRERQIHISERILSTYLGKSAEPVLPDLPLERLLDCNEDCGISGTGVGSPOL  
LVEPPTLESQTK"  
6051..6296  
/gene="vpu"  
6051..6296  
/codon\_start=1  
/product="vpu protein"  
/protein\_id="AAC28451.1"  
/db\_xref="GI:3378128"  
/translation="MOPIVAIIVAVIIIAVMSIVIIIEYRKILRQRIDRLDR  
LIERAEDSGNESEBSALIVEMGVEMHNAFWDIDL"  
6211..8775  
/gene="env"  
6211..8775  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AAC28452.1"  
/db\_xref="GI:3378129"  
/translation="MRYKERYOHLMRMGKWTMLGLILMISATEKILWTVYYGVY  
WKETTTLPFASDAPADTEYHNWATHACAPDPNPPEVYLVNVTNSHMTWDAYE  
OMHEDTISLMDOSLPCVKLPLCVSLKCTDILKNDTNTSSSGMTEKSGIKKCSFN  
IGTSIRKVOKEYAFYKLDIVPIDNTSYRLSCNTSYITQACPKVTFEPIPIHYCAP  
AGFAILCNKNTFNGTCTIVSTVQCTHGIRPVYSTOLLNGSLAEDVYIRSNFT"

gene  
CDS  
8777..9397  
/gene="nef"  
8777..9397  
/gene="nef"  
/codon\_start=1  
/product="nef protein"  
/protein\_id="AAC28453.1"  
/db\_xref="GI:3378130"  
/translation="MGKWSKGSYIGMPAVREBRRAEPADGVAVSRDLERHGLT  
SSNTAANNACAMLEAOEKEVEFPYQVPLRPMYKAADVLSHFLKEKGLGLIH  
SORODILDMITYTOGYFPDMONYTPGPVRYPLTGMCKYKIVPVPDREKANKGE  
NTSLHPSVLSHGMDDPERALEMFRDLAFHHVARELHREYFNRC"  
BASE COUNT 3417 a 1763 c 2355 g 2164 t  
ORIGIN  
Query Match 41.9%; Score 2607.4; DB 14; Length 9699;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 2635; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
DNAAKTIIVQNTSVEINCTPRYNNKRSIRIQRGPAFYTGKIGNMROAHCHISHA  
KWNATIKOIASKLEDEYTGNNKNTIIFKQSSGDEILKSNYANCPAGISQSSNTL  
WNTSTSEESNNTEGSDTITLPCRIQYIMMOCKYKAMVAPISQSSNTL  
LTPRDGNNNNNGSOIPRRGGGMDNRMSRLSYKYKIDLPICAPAKRKYVOREK  
RATGICLPIFGIGAGAGSTMGASMTITVYAROSLSGTYHOONLIRATAOHLIOL  
TWGIKOLQARILAVERYLKDQDLGLTWGSSGKICITVAVPMASNSKSLIEDIWNM  
TMMEMREINNTYSLIHSLESONOEBDELELDKNASLMMNFTNMWYIKL  
FIIVGGLVLRIVFAVVISVNRQOYSLSFQTHLPIRGDPRPGEIEGGERBR  
GRSIRLVNGSLALIMDLRSICLESYRRLDILLIYRIYELGRWELAKTWNNL  
OYMSQELKNSAVNLNLTATAVAEGRVIEVLOAVRAIRIRIIRIOLLENTLL"  
8777..9397  
/gene="nef"  
8777..9397  
/gene="nef"  
/codon\_start=1  
/product="nef protein"  
/protein\_id="AAC28453.1"  
/db\_xref="GI:3378130"  
/translation="MGKWSKGSYIGMPAVREBRRAEPADGVAVSRDLERHGLT  
SSNTAANNACAMLEAOEKEVEFPYQVPLRPMYKAADVLSHFLKEKGLGLIH  
SORODILDMITYTOGYFPDMONYTPGPVRYPLTGMCKYKIVPVPDREKANKGE  
NTSLHPSVLSHGMDDPERALEMFRDLAFHHVARELHREYFNRC"

QY	1951	GTTTTGGATTCTTAAATNGTAATTAATAAGCTTCATGTGAAACAGAGNCAATGTACAAATG	2010
Db	6869	GTTTTGGATTCTTAAATNGTAATTAATAAGCTTCATGTGAAACAGAGNCAATGTCAATTG	6928
QY	2011	TCACACAGATACAAATGTACACATAGAAATCAAGGCGAGTCAAGTCAACTCACTGCTGTAA	2070
Db	6929	TCACACAGATACAAATGTACACATAGAAATCAAGGCGAGTCAAGTCAACTCACTGCTGTAA	6988
QY	2071	ATGGCAGCTTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTACAGACAAATGCTA	2130
Db	6989	ATGGCAGCTTAGCAGAGAGAGATGTAGTAATTTAGATCTGCCAATTTACAGACAAATGCTA	7048
QY	2131	AAACCAATATGTACACTGAAACCACTGTAGAAATTAATGTACAAAGCCCAACAACA	2190
Db	7049	AAACCAATATATGACACTGAAACCACTGTAGAAATTAATGTACAAAGCCCAACAACA	7108
QY	2191	ATACAAAGAAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCAATTTGTATACATAGAA	2250
Db	7109	ATACAAAGAAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCAATTTGTATACATAGAA	7168
QY	2251	AAATAGCAAAATATGAGACAAGCACATTTGTACATTTAGTAGAGCAAAATGGAATGCCACTT	2310
Db	7169	AAATAGCAAAATATGAGACAAGCACATTTGTACATTTAGTAGAGCAAAATGGAATGCCACTT	7228
QY	2311	TAAACAGATATAGCTACAGCAAAATTAAGAGACAACTTTGGAATTAATTAATTAATCTTTA	2370
Db	7229	TAAACAGATATAGCTACAGCAAAATTAAGAGACAACTTTGGAATTAATTAATTAATCTTTA	7288
QY	2371	AGCAATCCTCAGAGGGGGAGCCACAGAAATTTGACGACAGTTTAAATTTGGAGGGGAAAT	2430
Db	7289	AGCAATCCTCAGAGGGGGAGCCACAGAAATTTGACGACAGTTTAAATTTGGAGGGGAAAT	7348
QY	2431	TTTTCTACTGTAAATTCACAACACACTGTTTAATAGTACTTGGTTTAATAGTACTTGAGTA	2490
Db	7349	TTTTCTACTGTAAATTCACAACACACTGTTTAATAGTACTTGGTTTAATAGTACTTGAGTA	7408
QY	2491	CTGAGAGGCTCAAAATTAACATCGAAGAGTAGACACAACTCACACTCCCATGCAAGATTAAC	2550
Db	7409	CTGAGAGGCTCAAAATTAACATCGAAGAGTAGACACAACTCACACTCCCATGCAAGATTAAC	7468
QY	2551	AATTTAATAACATGTGCGAGAAATGAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC	2610
Db	7469	AATTTAATAATATGTGGCAGAAATGAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC	7528
QY	2611	AAATTTAGATGTTTCATCAAAATATTACTGGCTGCTATTTAACAAAGATGTGTGTAAATACA	2670
Db	7529	AAATTTAGATGTTTCATCAAAATATTACTGGCTGCTATTTAACAAAGATGTGTGTAAATACA	7588
QY	2671	ACAATGGGTCGGAGATCTTCAGACCTGAGAGGGCGATATGAGGGGACAATTTGAGAAATG	2730
Db	7589	ACAATGGGTCGGAGATCTTCAGACCTGAGAGGGCGATATGAGGGGACAATTTGAGAAATG	7648
QY	2731	AATATATATTAATTAATAAGTATGTAATAATTTGACCATTAAGAGTAGACACCCACCAAGGCA	2790
Db	7649	AATATATATTAATTAATAAGTATGTAATAATTTGACCATTAAGAGTAGACACCCACCAAGGCA	7708
QY	2791	AGAGAAAGTGTGTCAGAGAGAAAAAAGAGAGTGGGAATTTGAGAGCTTTGCTTGGGT	2850
Db	7709	AGAGAAAGTGTGTCAGAGAGAAAAAAGAGAGTGGGAATTTGAGAGCTTTGCTTGGGT	7768
QY	2851	TCCTTGGAGCAGCAGGAAGCACTATGTGGGCTGACAGTCAATGACGCTAGCGGTACAGGCCA	2910
Db	7769	TCCTTGGAGCAGCAGGAAGCACTATGTGGGCGGCTCAATGACGCTAGCGGTACAGGCCA	7828
QY	2911	GACAAATATTTGTCTGATATATAGTGTGAGAGCAGAACAACTTTGTCTGAGGGCTATTAAGCGC	2970
Db	7829	GACAAATATTTGTCTGATATATAGTGTGAGAGCAGAACAACTTTGTCTGAGGGCTATTAAGCGC	7888
QY	2971	AACAGCAATGCTTCAACTCACACTGTGGGCGATCAAAACAGCTCCAGGCAAGATCTCTGG	3030
Db	7889	AACAGCAATCTGTCAACTCACACTGTGGGCGATCAAAACAGCTCCAGGCAAGATCTCTGG	7948
QY	3031	CTGTGGAAGATACCTAAGAGATACACAGCTCTGTGGGATTTGGGTTTGTCTGTGGAAC	3090

Db	7949	CTGTGGAAAGATACCTTAAGGATGACAGCTCCCTGGGGATTGGGTTGCTCTGGAAC	8008
QY	3091	TCATTTGGCACCACTGCTGTGCTTGGATGCTAGTTGGAGTAATAATCTCTGGACAGA	3150
Db	8009	TCATTTGGACACAGTGTGCTGCTTGGAAATGCTAGTTGGAGTAATAATCTCTGGACAGA	8068
QY	3151	TTTGGAAATACATGACTGCGATGGAGTGGGACAGAGAAATTTACAAATTTACCAAGCTTAA	3210
Db	8069	TTTGGAAATACATGACTGCGATGGAGTGGGAGAAATTTACAAATTTACCAAGCTTAA	8128
QY	3211	TACACTCCTTAATTGGAAGATCGGAAAAACGCAAGAAAGAAAGAACAAATTAATTGG	3270
Db	8129	TACACTCCTTAATTGGAAGATCGGAAAAACGCAAGAAAGAAAGAACAAATTAATTGG	8188
QY	3271	AATTAGATAAATGGGGCAAGTTTGTGGAATTGGTTTAACATTAACAAATTTGGCTGTGATA	3330
Db	8189	AATTAGATAAATGGGGCAAGTTTGTGGAATTGGTTTAACATTAACAAATTTGGCTGTGATA	8248
QY	3331	TAAATATTAATTCATTAATGATAGTAGAGGCTTGTAAGCTTTAAGAAATGTTTCTCTTAC	3390
Db	8249	TAAATATTAATTCATTAATTAATAGTAGAGGCTTGTAAGCTTTAAGAAATGTTTCTCTTAC	8308
QY	3391	TTTCTATATGTAATATAGATTTAGGAGGATTTTACCACTTATTCCTTTCAGACCACTCC	3450
Db	8309	TTTCTATATGTAATATAGATTTAGGAGGATTTTACATTAATTCCTTTCAGACCACTCC	8368
QY	3451	CAATCCGAGGGGAGCCGACGACGCGCCGAGAGAAATAGAAGAAAGTGGAGAGAGACA	3510
Db	8369	CAATCCGAGGGGAGCCGACGACGCGCCGAGAGAAATAGAAGAAAGTGGAGAGAGACA	8428
QY	3511	GAGACAGATCCATTCGATTAGTGAACGAGATCCTTAGCACTTATCTGGGACGATCTGGGA	3570
Db	8429	GAGCAGATCCATTCGATTAGTGAACGAGATCCTTAGCACTTATCTGGGACGATCTGGGA	8488
QY	3571	GCCGTGGCTCTTACGCTACCAACCGCTTGAAGACTTACCTTGAATGTAAAGAGAAFTG	3630
Db	8489	GCCGTGGCTCTTACGCTACCAACCGCTTGAAGACTTACCTTGAATGTAAAGAGAAFTG	8548
QY	3631	TGGAACCTTCTGGAGCGCAGGGGGTGGGAGGCCCTCAAAATATTGGTGGAAATCTCTACAGT	3690
Db	8549	TGGAACCTTCTGGAGCGCAGGGGGTGGGAGGCCCTCAAAATATTGGTGGAAATCTCTACAGT	8608
QY	3691	ATTGGAGTCAGGAACCTAAGAATATGTCGTGTTAACTTGCCTCAATGCCACAGCCTATACG	3750
Db	8609	ATTGGAGTCAGGAACCTAAGAATATGTCGTGTTAACTTGCCTCAATGCCACAGCCTATACG	8668
QY	3751	TAGCTGAGGGGACAGATAGGGTTTATAGAGTATTACAAAGCCTTATAGACTTATGGCC	3810
Db	8669	TAGCTGAGGGGACAGATAGGGTTTATAGAGTATTACAAAGCCTTATAGACTTATGGCC	8728
QY	3811	ACATATCCATGAAGAAATATAGACAGGGCTTGAAGAAAGATTTTCTTATAGATGGGGGACG	3870
Db	8729	ACATATCCATGAAGAAATATAGACAGGGCTTGAAGAAAGATTTTCTTATAGATGGGGGACG	8788
QY	3871	TGCTCAAAAATATAGTGTGATTGGATTTGGCTCTCTTAAAGGAGAAATAGACGAGCTGAG	3930
Db	8789	TGCTCAAAAATATAGTGTGATTGGATTTGGCTCTCTTAAAGGAGAAATAGACGAGCTGAG	8848
QY	3931	CCAGACAGATGGGGTGGGAGCAGATATCTCGAGATCTAGA 3971	
Db	8849	CCAGACAGATGGGGTGGGAGCAGATATCTCGAGACCTAGA 8889	
RESULT 10			
LOCUS	E00987	9213 bp RNA linear	PAT 29-SEP-1997
DEFINITION	Genomic DNA of HTLV-III encoding gag and p-24 protein.		
ACCESSION	E00987		
VERSION	E00987.1	GI:2169248	
KEYWORDS	JP 198623700-A/1.		
SOURCE	unidentified		
ORGANISM	unidentified		





Db 7181 ATACAAATGAGTCCGAGATCTTCAGACCTGGAGGAGAGATATGAGCAATTTGA 7240  
Oy 2725 GAATGAAATTTATTAATATATAAGTACGAAAAATTTGAACCATTTAGAGTACGACCCCA 2784  
Db 7241 GAAATGAAATTTATTAATATATAAGTACGAAAAATTTGAACCATTTAGAGTACGACCCCA 7300  
Oy 2785 AGGCAAGAGAGAGAGTGTGAGAGAGAAAAAGAGCGTGGSAATAGAGCTTTGTTC 2844  
Db 7301 AGGCAAGAGAGAGAGTGTGAGAGAGAAAAAGAGCGTGGSAATAGAGCTTTGTTC 7360  
Oy 2845 TTGGATCTTGGAGACAGACAGAGAACACTATGGCTGCACGCTCAATGACGCTGAC 2904  
Db 7361 TTGGATCTTGGAGACAGACAGAGAACACTATGGCGGACGCTCAATGACGCTGAC 7420  
Oy 2905 AGGCGACACATTTATGCTCATATAGTGCAGCAGCAGACAAATTTGCTGAGGCTATG 2964  
Db 7421 AGGCGACACATTTATGCTCATATAGTGCAGCAGCAGCAAAATTTGCTGAGGCTATG 7480  
Oy 2965 AGGCGCAACAGACATCTGTTGCAACGTCAGTGGGCGATCAACAGCTCCAGGCAAGAA 3024  
Db 7481 AGGCGCAACAGACATCTGTTGCAACGTCAGTGGGCGATCAACAGCTCCAGGCAAGAA 7540  
Oy 3025 TCTGGCTGTGAGAAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGCTTCTCTG 3084  
Db 7541 TCTGGCTGTGAGAAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGCTTCTCTG 7600  
Oy 3085 GAAACATCATTTGCACACCTGCTGCTTGGATGCTAGTTGGAGTATTAATCTCTGG 3144  
Db 7601 GAAACATCATTTGCACACCTGCTGCTTGGATGCTAGTTGGAGTATTAATCTCTGG 7660  
Oy 3145 AACAGATTTGGATTAACATGACGACGATGAGTGGAGAGAGAAATTTACAATTAACAA 3204  
Db 7661 AACAGATTTGGATTAACATGACGACGATGAGTGGAGAGAGAAATTTACAATTAACAA 7720  
Oy 3205 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATGACAGAAAT 3264  
Db 7721 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATGACAGAAAT 7780  
Oy 3265 TATTGGAATTAATTAATGAGCAAGTTGTTGGAATTTGTTTAACATAACAAATTTGCTGT 3324  
Db 7781 TATTGGAATTAATTAATGAGCAAGTTGTTGGAATTTGTTTAACATAACAAATTTGCTGT 7840  
Oy 3325 GGTATATTAATTTATTCATATAGTAGAGAGGCTTGTAGGTTTAAGATAGTTTGG 3384  
Db 7841 GGTATATTAATTTATTCATATAGTAGAGAGGCTTGTAGGTTTAAGATAGTTTGG 7900  
Oy 3385 CTGACTTTCTATAGTAATAGAGTATAGGAGGATATTCACATTTATGTTTACAGACC 3444  
Db 7901 CTGACTTTCTATAGTAATAGAGTATAGGAGGATATTCACATTTATGTTTACAGACC 7960  
Oy 3445 ACCTCCCAATCCGAGGAGGAGCCGAGAGCCGAGAGAAATAGAAAGAGAGCTGGAGAGA 3504  
Db 7961 ACCTCCCAATCCGAGGAGGAGCCGAGAGCCGAGAGAAATAGAAAGAGAGCTGGAGAGA 8020  
Oy 3505 GAGACAGAGACAGATCCATTCGATTAAGTAAGAGGATTCCTTACACTTTCTGGGAGCTTC 3564  
Db 8021 GAGACAGAGACAGATCCATTCGATTAAGTAAGAGGATTCCTTACACTTTCTGGGAGCTTC 8080  
Oy 3565 TGGGAGGCTGTGCTCTTACAGTACCAACCGCTTGAAGACTTACTTTGATTTGAAGA 3624  
Db 8081 TGGGAGGCTGTGCTCTTACAGTACCAACCGCTTGAAGACTTACTTTGATTTGAAGA 8140  
Oy 3625 GGATTTTGGAACTTCTGGAGAGCAGAGGGGTGGAAAGCCCTCAATATTTGGTGAATCTCC 3684  
Db 8141 GGATTTTGGAACTTCTGGAGAGCAGAGGGGTGGAAAGCCCTCAATATTTGGTGAATCTCC 8200  
Oy 3685 TACAGTATTTGAGTACAGAACTTAAGAATAGTGGCTTAACCTGTCATGCTCCAGAGCA 3744  
Db 8201 TACAGTATTTGAGTACAGAACTTAAGAATAGTGGCTTAACCTGTCATGCTCCAGAGCTA 8260  
Oy 3745 TAGCAGTAGCTGAGGGAGACAGATAGGCTTATAGAAGTATTAACAAGCACTTATAGAGCTA 3804  
Db 8261 TAGCAGTAGCTGAGGGAGACAGATAGGCTTATAGAAGTATTAACAAGCACTTATAGAGCTA 8320

Oy 3805 TTCGCCACATACCTAGAGAAATTAAGACAGAGGCTTGGAAAGATTTTCTATAGATGGGT 3864  
Db 8321 TTCGCCACATACCTAGAGAAATTAAGACAGAGGCTTGGAAAGATTTTCTATAGATGGGT 8380  
Oy 3865 GGCAAGTGTGTAAAAAGAGTGTGATTTGAGAGGCTGCTGTAAAGGAAGAATGAGACGA 3924  
Db 8381 GGCAAGTGTGTAAAAAGAGTGTGATTTGAGAGGCTGCTGTGTAAAGGAAGAATGAGACGA 8440  
Oy 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGACAGTATCTGAGATCTAGA 3971  
Db 8441 GCTGAGCCAGCAGCAGATGGGGTGGAGACAGTATCTGAGACCTAGA 8487  
RESULT 11  
104549 104549 9213 bp DNA linear PAT 02-DEC-1994  
LOCUS Sequence 11 from Patent EP 0187041.  
DEFINITION 104549  
ACCESSION 104549  
VERSION 104549.1 GI:591641  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9213)  
AUTHORS Capon,D.J. and Lasky,L.A.  
TITLE Molecularly cloned acquired immunodeficiency syndrome polypeptides,  
intermediates therefor and methods and materials for their use  
JOURNAL Patent: EP 0187041-A1 11 09-JUL-1986;  
FEATURES  
source location/Qualifiers  
1..9213  
BASE COUNT 3297 a 1655 c 2217 g 2044 t  
ORIGIN  
Query Match 41.7%; Score 2599.4; DB 6; Length 9213;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 2640; Conservative 0; Mismatches 41; Indels 6; Gaps 1;  
Oy 1291 CCATGAGAGTAGAGAGAGATATCAGCACTTGTGAGATGGGGGTGAATGGGCAACA 1350  
Db 5802 CAATGAGAGTAGAGAGAGAAATATCAGCACTTGTGAGATGGGGGTGAATGGGCAACA 5861  
Oy 1351 TGCCTCTGGGATATGATGATCTGTAGCTTACAGAAATTTGGGTCAACGCTATT 1410  
Db 5862 TGCCTCTGGGATATGATGATCTGTAGCTTACAGAAATTTGGGTCAACGCTATT 5921  
Oy 1411 ATGGGATCTGTGTGGAAGAGCAACACCACTATTTTGTGATCAGATGCTAAAG 1470  
Db 5922 ATGGGATCTGTGTGGAAGAGCAACACCACTATTTTGTGATCAGATGCTAAAG 5981  
Oy 1471 CATATGATACAGAGGTACATATGTTTGGGCAACATGCTGTGTACCAAGACCCA 1530  
Db 5982 CATATGATACAGAGGTACATATGTTTGGGCAACATGCTGTGTACCAAGACCCA 6041  
Oy 1531 ACCCAAGAAGTAGTATTTGTAATGTGACAGAAATTTTAACTATGGAATAATGCA 1590  
Db 6042 ACCCAAGAAGTAGTATTTGTAATGTGACAGAAATTTTAACTATGGAATAATGCA 6101  
Oy 1591 TGTGAGAACAGATGATGATTAATCACTTTATGGATCAAAAGCTTAAAGCATGTG 1650  
Db 6102 TGTGAGAACAGATGATGATTAATCACTTTATGGATCAAAAGCTTAAAGCATGTG 6161  
Oy 1651 TAAATTAACCCACTGTGTAGTTTAAAGTCACTGATTTGAAGATATACTAATA 1710  
Db 6162 TAAATTAACCCACTGTGTAGTTTAAAGTCACTGATTTGAAGATATACTAATA 6221  
Oy 1711 CCAATAGTAGAGGGGAGATTAATAGGAGAAAGAGATATAAAATCTGCTTTCA 1770  
Db 6222 CCAATAGTAGAGGGGAGATTAATAGGAGAAAGAGATATAAAATCTGCTTTCA 6281  
Oy 1771 ATATCAGCACAAGCATAAGATTAAGTGAAGAAAGATATGCAATTTTATTAACCTTG 1830  
Db 1771 ATATCAGCACAAGCATAAGATTAAGTGAAGAAAGATATGCAATTTTATTAACCTTG 1830

Dh 6282 ATATCAGACAGACATGAAGGTAAAGGTGCAGAAAATAATGCAATTTTTTATTAACCTTG 6341  
Oy 1831 ATATAGTACCAATGATATA-----TACCAGCTTAGTGTGATAGTTGTTAAACCTCAG 1884  
Dh 6342 ATATATATACCAATGATATGATATCTACCACTTATAGCTTGACAAAGTTGTAACCTCAG 6401  
Oy 1885 TCATTTACAGAGCCCTGTCAAAGAGTATCCTTGAAGCAATTCCTCATATATTTGTCCC 1944  
Dh 6402 TCATTTACAGAGCCCTGTCAAAGAGTATCCTTGAAGCAATTCCTCATATATTTGTCCC 6461  
Oy 1945 CCGCTGGTTTTGGCATCTTAAATGTAATTAAGTCTCAATGAAGAAAGGACCAATGTA 2004  
Dh 6462 CCGCTGGTTTTGGCATCTTAAATGTAATTAAGTCTCAATGAAGAAAGGACCAATGTA 6521  
Oy 2005 CAATGTGCAGCAGCAGTACAAATGTACATGGAATCAGGCCAGTAGTATCAACTCACTGTC 2064  
Dh 6522 CAATGTGCAGCAGCAGTACAAATGTACATGGAATTAGGCCAGTAGTATCAACTCACTGTC 6581  
Oy 2065 TGTTTAAATGGCAGCTCTAGCAGAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCAGACA 2124  
Dh 6582 TGTTTAAATGGCAGCTCTAGCAGAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCAGACA 6641  
Oy 2125 ATGCTTAAACCATTAATGTACAGTGAACATCTGTAGAATTAATTTGTAACAAGCCA 2184  
Dh 6642 ATGCTTAAACCATTAATGTAGTACAGTGAACATCTGTAGAATTAATTTGTAACAAGCCA 6701  
Oy 2185 ACAACAATAGCAGAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGAGCAATTTGTTACAA 2244  
Dh 6702 ACAACAATAGCAGAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGAGCAATTTGTTACAA 6761  
Oy 2245 TAGGAAAAATAGGAAAAATATGAGACAGACACTTTGTAACATTTAGTAGACCAAAATGGAATG 2304  
Dh 6762 TAGGAAAAATAGGAAAAATATGAGACAGACACTTTGTAACATTTAGTAGAGCAAAATGGAATA 6821  
Oy 2305 CCACCTTAAAAACAGATGCTAGCAGAAATTAAGAGAACATTTGGAATTAATTAACAATTA 2364  
Dh 6822 ACACCTTAAAAACAGATGATATGACAAATTAAGAGAACATTTGGAATTAATTAACAATTA 6881  
Oy 2365 TCTTTAAGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAACGACACAGTTTAAATTTGGAG 2424  
Dh 6882 TCTTTAAGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAACGACACAGTTTAAATTTGGAG 6941  
Oy 2425 GGGAAATTTTCTACTGTATTAATCAACACACTGTTTAATAGTACTTGGTTTAAATGTAAT 2484  
Dh 6942 GGGAAATTTTCTACTGTATTAATCAACACACTGTTTAATAGTACTTGGTTTAAATGTAAT 7001  
Oy 2485 GGAATGCTGAGAGGCTCAATTAACACTGAGAGAGTGAACACATACACTCCCATGCAAGAA 2544  
Dh 7002 GGAATGCTGAGAGGCTCAATTAACACTGAGAGAGTGAACACATACACTCCCATGCAAGAA 7061  
Oy 2545 TAAACAATTTTATTAACATGTGCGAGAGAGTAGGAAAAAGCAATGTATGCCCTCCATCA 2604  
Dh 7062 TAAACAATTTTATTAACATGTGCGAGAGAGTAGGAAAAAGCAATGTATGCCCTCCATCA 7121  
Oy 2605 GTGAGCAAAATTTAGATGTTCAATCAAAATTTACTGGGCTGTATTAACAAGAGATGGTGTGA 2664  
Dh 7122 GCGGACAAATTTAGATGTTCAATCAAAATTTACTAGGGGCTGTATTAACAAGAGATGGTGTGA 7181  
Oy 2665 ATTAACAATTTGGTCCGAGATCTTCAGACCTGAGAGAGCGATATAGAGCAAAATTTGGA 2724  
Dh 7182 ATTAACAATTTGAGTCCGAGATCTTCAGACCTGAGAGAGCGATATAGAGCAAAATTTGGA 7241  
Oy 2725 GAAGTGAATTTATATTAATTAAGTATTAATAAATTTGAACCAATTTGAGAGTGAACCCAGCA 2784  
Dh 7242 GAAGTGAATTTATATTAATTAAGTATTAATAAATTTGAACCAATTTGAGAGTGAACCCAGCA 7301  
Oy 2785 AGGCAAGAGAGAGAGTGTGCAGAGAGAAAAAGAGCAAGTGGGAATAGAGCTTTGTCTCC 2844  
Dh 7302 AGGCAAGAGAGAGAGTGTGCAGAGAGAAAAAGAGCAAGTGGGAATAGAGCTTTGTCTCC 7361  
Oy 2845 TTGGGTTCTTTGGAGCAGCAGAGAAAGCACTATGGGCTCAGCTCATATAGCGTGAAGCTAC 2904  
Dh 7362 TTGGGTTCTTTGGAGCAGCAGAGAAAGCACTATGGGCTCAGCGTCAATATAGCGTGAAGCTAC 7421

Oy 2905 AGGCAAGACAAATTTATGCTGATTAAGTGAAGCAGACAGAAACAATTTGCTGAGGGCTATTTG 2964  
Dh 7422 AGGCAAGACAAATTTATGCTGATTAAGTGAAGCAGACAGAAACAATTTGCTGAGGGCTATTTG 7481  
Oy 2965 AGGCGCAACAGCATCTGTTCCACTACACTCAGCTGTGGGCTATCAACAGCTCCAGGCAAGAA 3024  
Dh 7482 AGGCGCAACAGCATCTGTTCCACTACACTCAGCTGTGGGCTATCAACAGCTCCAGGCAAGAA 7541  
Oy 3025 TCCGAGCTGTGGAAAGATTAAGTAAAGATCAACAGCTGTGGGCTATTTGGGTTGCTGCTG 3084  
Dh 7542 TCCGAGCTGTGGAAAGATTAAGTAAAGATCAACAGCTGTGGGCTATTTGGGTTGCTGCTG 7601  
Oy 3085 GAAACATCAATTTGCACACCTGTGCTGCTTGAATGCTAGTTGGAGTAAATTAATCTCTGG 3144  
Dh 7602 GAAACATCAATTTGCACACCTGTGCTGCTTGAATGCTAGTTGGAGTAAATTAATCTCTGG 7661  
Oy 3145 AACGATTTTGGAAATTAACATGACCTGGATGAGTGGGACAGAGAAATTAACAAATTTACCAA 3204  
Dh 7662 AACGATTTTGGAAATTAACATGACCTGGATGAGTGGGACAGAGAAATTAACAAATTTACCAA 7721  
Oy 3205 GCTTAATACACTCCTTAATTTGAAGATTCGCAAAACAGCAAGAAAGAAATGAACAAAGAT 3264  
Dh 7722 GCTTAATACACTCCTTAATTTGAAGATTCGCAAAACAGCAAGAAAGAAATGAACAAAGAT 7781  
Oy 3265 TATTTGAATTAAGATTAATGGGCAAGTTTGTGAATTTGTTAAACATTAACAATTTGCTGT 3324  
Dh 7782 TATTTGAATTAAGATTAATGGGCAAGTTTGTGAATTTGTTAAACATTAACAATTTGCTGT 7841  
Oy 3325 GGTATATAAATTAATTTATTAATGATAGTGAAGGCTGTGTAAGTTTAAAGTAATTTGTTG 3384  
Dh 7842 GGTATATAAATTAATTTATTAATGATAGTGAAGGCTGTGTAAGTTTAAAGTAATTTGTTG 7901  
Oy 3385 CTGTACTTTCTATGTAATTAAGATTAAGTGAAGGATTTTACATTAATTTGCTTCAGACCC 3444  
Dh 7902 CTGTACTTTCTATGTAATTAAGATTAAGTGAAGGATTTTACATTAATTTGCTTCAGACCC 7961  
Oy 3445 ACCTCCCAATCCGAGGGGACCCGACAGGCCCGAAGAAATGAAGAAAGTTGGAGAGA 3504  
Dh 7962 ACCTCCCAATCCGAGGGGACCCGACAGGCCCGAAGAAATGAAGAAAGTTGGAGAGA 8021  
Oy 3505 GAGACAGAGACATTCATTCATGATTAAGAACGATTCCTTAAGCACTTATCGGAGCATC 3564  
Dh 8022 GAGACAGAGACATTCATTCATGATTAAGAACGATTCCTTAAGCACTTATCGGAGCATC 8081  
Oy 3565 TGCAGAGCCGTGCTGCTTTCAGTACCAAGCTTGAGAGACTTACTCTTATTTGTAAGCA 3624  
Dh 8082 TGCAGAGCCGTGCTGCTTTCAGTACCAAGCTTGAGAGACTTACTCTTATTTGTAAGCA 8141  
Oy 3625 GGATTTGTGAACCTTCTGGAGCGAGGGGTGGGAAGCCCTCAATTAATTTGTGAATCTCC 3684  
Dh 8142 GGATTTGTGAACCTTCTGGAGCGAGGGGTGGGAAGCCCTCAATTAATTTGTGAATCTCC 8201  
Oy 3685 TACAGTATTTGGAGTCAAGAACTTAAGAAATGCTGTTAACCTGTCTCAATGCCACAGCA 3744  
Dh 8202 TACAGTATTTGGAGTCAAGAACTTAAGAAATGCTGTTAACCTGTCTCAATGCCACAGCA 8261  
Oy 3745 TAGCAGTAGCGAGGGGACAGATAGGTTTGAAGATTAACAAGCAAGCTTATAGACCTA 3804  
Dh 8262 TAGCAGTAGCGAGGGGACAGATAGGTTTGAAGATTAACAAGCAAGCTTATAGACCTA 8321  
Oy 3805 TTGCGCCACATTAACCTTAAGAAATTAAGACAGAGCTTGGAAGGATTTTCTTAAGATGGGT 3864  
Dh 8322 TTGCGCCACATTAACCTTAAGAAATTAAGACAGAGCTTGGAAGGATTTTCTTAAGATGGGT 8381  
Oy 3865 GGCAGATGTGTCAAAAAGTACTGTGATTTGATGCGCTGCTGTAAAGGAAGATGACAGCA 3924  
Dh 8382 GGCAGATGTGTCAAAAAGTACTGTGATTTGATGCGCTGCTGTAAAGGAAGATGACAGCA 8441  
Oy 3925 GCTGAGCCACAGCAGATTTGGGTTGGGAGCAGTATCTGGAATCTGAGA 3971  
Dh 8442 GCTGAGCCACAGCAGATTTGGGTTGGGAGCAGCAGTATCTGGAACCTGAGA 8488

RESULT 12  
HYMCK1  
LOCUS 9752 bp DNA linear VRL 25-DEC-2002  
DEFINITION Human immunodeficiency virus 1 DNA for Gag, Pol, Vif, Vpr, Tat,  
Rev, Env, Vpu.  
ACCESSION DB6068.1 GI:1398973  
VERSION DB6068.1  
KEYWORDS Vpu; Env; Rev; Tat; Vpr; Vif; Pol; Gag.  
SOURCE Human immunodeficiency virus 1  
ORGANISM Human immunodeficiency virus 1  
Virus; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 9752)  
Cloyd,M.W. and Moore,B.E.  
Spectrum of biological properties of human immunodeficiency virus  
TITLE Virology 174 (1), 103-116 (1990)  
JOURNAL  
MEDLINE 90101366  
PUBMED 1688473  
REFERENCE 2 (bases 1 to 9752)  
Iwatahi,Y.  
Unpublished  
REFERENCE 3 (bases 1 to 9752)  
Iwatahi,Y.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (17-JUN-1996) Yasunasa Iwatahi, Institute for Virus  
Research, Kyoto University, Laboratory of Gene Analysis; Sakyo-ku,  
Shogoin-Kawahara-machi, Kyoto, Kyoto 606-01, Japan  
(E-mail:yiwatahi@virus.kyoto-u.ac.jp, Tel:075-751-3996,  
Fax:075-751-3995)  
FEATURES  
source  
1..9752  
location/Qualifiers  
/organism="Human immunodeficiency virus 1"  
/mol\_type="genomic DNA"  
/strain="MCK1"  
/isolate="MCK"  
/db\_xref="taxon:11676"  
1..636  
790..2328  
/gene="gag"  
790..2328  
/gene="gag"  
/codon\_start=1  
/product="gag"  
/protein\_id="BA12988.1"  
/db\_xref="GI:1398974"  
/translation="MGARASYLGGELDRMEKIRLRPGKKKKYKLIHWASRELER  
AVNGLLETSGCROILGOLQPSLOTGSEELRSLYNTVATLYCHORLEIKDTEALD  
KIEEONKSKKKAQAAADGSHSSOVSONPIYONIGOWHOALISPTLNAMKVVE  
EKASPEVIRPESALSEGATPODNTMNTYGGQAAMOMLKETINEAAMWDVHYV  
HAGPIAFGOMREPRGSDIAGTSTLQDIGMTNPPIVEITKRWLITGLINKIVRA  
YSPSTILDIRGPKPEPRDYDRYKTLRAEQASQEVKNMNTETLIVONAPDCITIL  
KALGPAATLEMTACOGVGGPGHARVLAEMQOVYNSATIMQRGFRNQRKIVLC  
FNCGKEGHIANRCAPRRKCGMCKEGHOKHDCQERANFLGKIMPSYKGRPNFLQ  
SRPEPTAPPLFQSRPEPTAPPESEFRSGVETTTSGQEPIDKELYPLTSLRSLFND  
PSQO"  
2085..5132  
/gene="pol"  
/gene="pol"  
/gene="pol"  
/codon\_start=1  
/product="pol"  
/protein\_id="BA12989.1"  
/db\_xref="GI:1398975"  
/translation="FFREDLAFLOGKAREFSEQOTRANSPITSEQOTRANSPTRRELQ  
VMGRDNNSLSEAGAROGTVSFNFPOITLMORPLVTIKIGOLKEALLDAGDVTLE  
EMSLBGRMKPKMIGIGIGELIKROYDILIEIGHKAITVLGTPPNILIGRLTLT  
IGCTLNPISPITGIVPVKLRGMDGPKYKOPULDEBKIKALVEICTEKKEKSKISKIS  
PENNYNPVPAIKKSDTKWRKLVDFRELNRQDDEFWOLGIPHAGLKKRKSATVL  
DVGDAYFSPLEDEPRKYTATTPISINNEPGIYOYNVLPOGKGSFALPQSSMTKI  
LEPKKONPDIVIOYMDLIVGSDLEIGORHRIEELRQHLRLRMGFTTPDKKQKEP  
PLTMMGYELHBDKWTQPIVLPEKDSMTVNDIQKLVGKLMAWSQIYPGIKVROCKLIL

RGTKALFEVILPTEAELELAENREILKEPHGVYYPDPSKDLIAEIOKOGOGWTOYI  
YOBEPKILTKGVAMREGAHNNDYKOLTELEAOKITTESIYVWCKTPKPKLIDKETME  
TWTEYMOATIPPEKEPVNTPPLVYKALOLEKEPVEVGAETFYDGAANRETRLRKAGI  
VTNRGQKQVPLTNTTQKTELQAIYIALQDSGLEVNLVDSOYALGIIQAPQOSSES  
ELVNDQILEOLIKKEKRYLAWPAHKGISGNDQVRLVSAIGIRKILFDGIDKADDEHE  
KLYNNMRMASDPNLPVYAKEIVASCDQOLKEAMGQVDCSPGILQDLCTHLEK  
VILVAVHASGVTIAEYVAPTEGOETVFEILKIGRMRVKTILHNDGNSPSTYKKA  
CMAKICQEGEIPYNNPOSQGVSEMMKEIKIIGOVROAEHLKTAQVMAVFTNPKR  
KRIGIGISAGERIYDITATDIQTELOKQITKTONFRYTYDSRNPIMKGAKLIMKG  
EGAVVINDNSDIKVPRRKAKIIRDYGMAGDQCVASRODED"  
5077..5655  
/gene="vif"  
5077..5655  
/gene="vif"  
/gene="vif"  
/codon\_start=1  
/product="vif"  
/protein\_id="BA12990.1"  
/db\_xref="GI:1398976"  
/translation="MENRWQVMIYQVDRMRIRTMKSLVKHHMYVSGKARGPYRHHY  
ESPHRISSEVHILGDARLIYTYMKILHGERDMHIGQYSIEMKRRKYTOVDREL  
ADQILHLTFDCSDATRKALGLHISPRCEYQGHKVGSLQTLALALITRKIK  
PPLPSVTKLTEDRNKKPQTKGHRGSHTMNGH"  
5595..5708  
/gene="vpr"  
5595..5708  
/gene="vpr"  
/codon\_start=1  
/product="vpr"  
/protein\_id="BA12991.1"  
/db\_xref="GI:1398977"  
/translation="MEQAPEDQGPQREPHNEMTLELELKNEAVRHPRP"  
join(5867..6081,8412..8457)  
/gene="tat"  
join(5867..6081,8412..8457)  
/gene="tat"  
/codon\_start=1  
/product="tat"  
/protein\_id="BA12992.1"  
/db\_xref="GI:1398978"  
/translation="MEPYDPRLEPMKHPGSGPKTACTNCYCKKCCFHCQVCFITKALG  
ISYGRKKRRORRRPQGSQTHQVSLSKQPTSQPRQDPGPKE"  
join(6006..6081,8412..8686)  
/gene="rev"  
join(6006..6081,8412..8686)  
/gene="rev"  
/codon\_start=1  
/product="rev"  
/protein\_id="BA12993.1"  
/db\_xref="GI:1398979"  
/translation="MAGSGSDSDLLKAVRLIKELYOSNPPRPNDEGTROARRRRR  
WRERQIHSTISERILSTYLRGRSADAVPLQPLERLITLDCDEDCGTSGTGCVSPQI  
LVESPTLIESGAK"  
6098..6343  
/gene="vpu"  
6098..6343  
/gene="vpu"  
/codon\_start=1  
/product="vpu"  
/protein\_id="BA12994.1"  
/db\_xref="GI:1398980"  
/translation="MOPIDQIALVALVAIIIAIVWASIVILEYRKLRLQRKIDRLIDR  
LIERAEDSGNSEKISALVEMGVEMGHNAWDDLL"  
6258..8828  
/gene="env"  
6258..8828  
/gene="env"  
/codon\_start=1  
/product="env"  
/protein\_id="BA12995.1"  
/db\_xref="GI:1398981"  
/translation="MRVKEKYOHLMRMGRWGTMLGLMLICATSEKLMVTVYGVVPY  
WKEAATTLTFCSDAKKAVDEWHNVAWIAHACVPIDPNQEVVLVAVTTEFNMKKMDVE  
QMHDITISLMDQSLKPCVKLPLCVLSIKCTDLKKNDJTNWSSSGRMIMKGBIKKCSFN

APMSEKREKVOEAFAPKEDJIDINDSTSYLSRVSNTVQAKXSPERIPJHC  
 APAPAFALCKNKNTFNGTQPCNTVSTYOTCEITREPVSTQTLNLSABEVAIRSN  
 FTDNAKJITVOLNOVEINOTPRNNTRKIRIQPGGAFTYTGKJNMOAHCNS  
 RAKMANALIKJASKEAQEENNTTIFKQSGDGEVTHASMGPEYFQNSJNT  
 STWNSWMTSGESINTEGJITLPCRIKQFINMOEYGMKAPISQJICRSLT  
 GILLTRPGCKNNSESEFPGGCDMDNRSLEYKVKJLEZVNLAPTRAKRVOY  
 EKRAVJGALFELEFGAAGCTSMJTLVQARQJSEYOONNLATRAEAOH  
 OJTWYMKOLARJIAVERIKKQOJLIGWCSGKLTCTAPMNAKSNSTNMA  
 NEMMBEDRENTNYSJLSIEESENQOEKREKEDKASJNNENTNMAWT  
 KJPMIYGLVGRIPYAVUSVYNRKQSTSPJQTHLPTRPGDRBEGIEEBGR  
 DRDSJISVMSALJIMDRTJCYSHRLDILITVRIJVEILGRGKALMYNNAL  
 LLQVWSOEKNSVSLNMTAIAVAGCTREVEVQACARJIRHPJIRGOREJLE

repeat_region	9117	.9752		
BASE COUNT	3432 a	1779 c	2368 g	2173 t
ORIGIN				

Query Match	41.7%;	Score 2539,4;	DB 14;	Length 9752;
Best Local Similarity	98.3%;	Pred. NO. 0;		
Matches 2640;	Conservative 0;	Mismatches 41;	Indels 6;	Gaps 1;

OY	1291	CCATGAGACGTCACAGACAGTATACAGACTCTGTGACATGAGGGGTGGAAATGGGGACACA	1350
Db	6256	CAATGAGAGTCGAAGAAAAATATCAGACTTGTGGAATGGGGGTGGAGATGGGGACACA	6315
OY	1351	TGCTCTTTGGGATTTTATGATCTGTATGTCGTACAGAAAAATTTGGGTACACCTCTATT	1410
Db	6316	TGCTCTTTGGGATTTTATGATCTGTATGTCGTACAGAAAAATTTGGGTACACCTCTATT	6375
OY	1411	ATGGGGTACCTGTGTGGAAGGAAAGCAACACCACTCTATTTTTGCATCAGATGCTTAAG	1470
Db	6376	ATGGGGTACCTGTGTGGAAGGAAAGCAACACCACTCTATTTTTGCATCAGATGCTTAAG	6435
OY	1471	CATTTGATACAGAGGTCACATATATTTTGGGACACATTCGCTGTGTACCAAGACCCCA	1530
Db	6436	CATTTGATACAGAGGTCACATATATTTTGGGACACATTCGCTGTGTACCAAGACCCCA	6495
OY	1531	ACCCACAAGATGTGTGTTGGTAAATGTCAGAAAAATTTTAACATGTGGAAAAATGACA	1590
Db	6496	ACCCACAAGATGTGTGTTGGTAAATGTCAGAAAAATTTTAACATGTGGAAAAATGACA	6555
OY	1591	TGTGAGAACAGATCATGAGGATTTAATCAGTTTATGGGATCAAAAGCTTAAAGCCATGTG	1650
Db	6556	TGTGAGAACAGATCATGAGGATTTAATCAGTTTATGGGATCAAAAGCTTAAAGCCATGTG	6615
OY	1651	TAAATTTAACCCACTCTGTGTAGTTTAAAGTGCACCTGATTTGAGAAATGATCTATA	1710
Db	6616	TAAATTTAACCCACTCTGTGTAGTTTAAAGTGCACCTGATTTAAGAAATGATCTATA	6675
OY	1711	CCATAGTAGTAGCGGGAGATGTAAATGGAGAAAGAGAGATAAAAAAGCTCTTTTCA	1770
Db	6676	CCATAGTAGTAGCGGGAGATGTAAATGGAGAAAGAGATAAAAAAGCTCTTTTCA	6735
OY	1771	ATATCACACACAGATTAAGATTAAGTGGAGAAAGAAATATGCATCTCTTTAATAACTTG	1830
Db	6736	ATATCACACACAGATTAAGATTAAGTGGAGAAAGAAATATGCATCTCTTTAATAACTTG	6795
OY	1831	ATATAGTACCAATATGATATA-----TACAGACTTATAGTTGATTAAGTTGAACACCTCAG	1884
Db	6796	ATATAGTACCAATATGATATACTACAGACTTACGTTGACAAAGTTGTGAACCTCAG	6855
OY	1885	TCATTTACAGAGGCTGTCCAAAGGTATCCTTTGAGCCAAATTCOCATCATTTATGTGCC	1944
Db	6856	TCATTTACAGAGGCTGTCCAAAGGTATCCTTTGAGCCAAATTCOCATCATTTATGTGCC	6915
OY	1945	CGGCTGTTTTGCGATTTCTAAATGTATATAAAGACGTTTCATATGAAACAGACCATGTA	2004
Db	6916	CGGCTGTTTTGCGATTTCTAAATGTATATAAAGACGTTTCATATGAAACAGACCATGTA	6975
OY	2005	CAATATGCACAGCATGTACATGTACACATATGGAATCAGGCCAGTAGTATCAATCTAATGC	2064
Db	6976	CAATATGCACAGCATGTACATGTACACATATGGAATATGAGCAAGTAGTATCAATCTAATGC	7035

QY	2065	TGTTAAATGCGAGCTCTAGCAGAAAGAAAGATGTAGTAATTAATGATCTCGCCAAATTTTCAGACACA	2124
Db	7036	TGTTAAAGGAGAGCTCTAGCAGAAAGAGAGTAGTAATTAATGATCTCGCCAAATTTTCAGACACA	7095
QY	2125	ATGCTAAAGACCATTAATAGTACAGCTGGAACACATCTGTAGAAAATTTAAATTTGTAACAAGACCCA	2184
Db	7096	ATGCTAAAGACCATTAATAGTACAGCTGGAACACATCTGTAGAAAATTTAAATTTGTAACAAGACCCA	7155
QY	2185	ACAACAATATCAAGAAAAAGTATCCGTAATCCAGAGGGGACGAGGGAGAGCATTTGTTCACAA	2244
Db	7156	ACAACAATATCAAGAAAAAGTATCCGTAATCCAGAGGGGACGAGGGAGAGCATTTGTTCACAA	7215
QY	2245	TAGGAAAAAATAGGAATATGTAGACACACACTTTGAACATTATAGTAGAGCAAAATGGATG	2304
Db	7216	TAGGAAAAAATAGGAATATGTAGACACACACTTTGAACATTATAGTAGAGCAAAATGGATG	7275
QY	2305	CCACTTTAAACAGATGAGCTAGCAAAATTAACAGAACATTTGGAAATTAATTAACAACTAA	2364
Db	7216	CCACTTTAAACAGATGAGCTAGCAAAATTAACAGAACATTTGGAAATTAATTAACAACTAA	7335
QY	2365	TCTTTTAAGCAATCCCTCAGAGAGGGGACCCAGAAATTTGTACGCACACAGTTTAAATTTGTGGAG	2424
Db	7336	TCTTTTAAGCAATCCCTCAGAGAGGGGACCCAGAAATTTGTACGCACACAGTTTAAATTTGTGGAG	7395
QY	2425	GGGAATTTTCTACTGTAATTCACACACACTGTTTAAATAGTACTTGTATTAATAGTACTT	2484
Db	7396	GGGAATTTTCTACTGTAATTCACACACACTGTTTAAATAGTACTTGTATTAATAGTACTT	7455
QY	2485	GGAGTACTGAAGGGTCAAATTAACCTGAGGAAGTGTACACAACTACACCTCCCATGCGAA	2544
Db	7456	GGAGTACTGAAGGGTCAAATTAACCTGAGGAAGTGTACACAACTACACCTCCCATGCGAA	7515
QY	2545	TAAACAATTTATTAACATGTGGCAGAGAAAGTAGAGAAAGCAATGTATAGCCCTCCCATCA	2604
Db	7516	TAAACAATTTATTAACATGTGGCAGAGAAAGTAGAGAAAGCAATGTATAGCCCTCCCATCA	7575
QY	2605	GTGGACAATTAAGATGTTTCATCAAAATTAATTACTGGGCTGTATTAATACAGAGATGTGTGA	2664
Db	7576	GTGGACAATTAAGATGTTTCATCAAAATTAATTACTGGGCTGTATTAATACAGAGATGTGTGA	7635
QY	2665	ATTAACAACAATGGGTCGAGATCTTTCAGAGCTGGAGAGAGCGGATATAGAGGACAAATTTGGA	2724
Db	7636	ATTAACAACAATGGGTCGAGATCTTTCAGAGCTGGAGAGAGCGGATATAGAGGACAAATTTGGA	7695
QY	2725	GAAGTGAATTTATTAATATTAATTAAGTAGTAAGAAATTTGAACCAATTAGAGTAGACACCCACCA	2784
Db	7696	GAAGTGAATTTATTAATATTAATTAAGTAGTAAGAAATTTGAACCAATTAGAGTAGACACCCACCA	7755
QY	2785	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTCTTC	2844
Db	7756	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTCTTC	7815
QY	2845	TTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCTGCAGTCAATATACCTGTACGGTAC	2904
Db	7816	TTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCTGCAGTCAATATACCTGTACGGTAC	7875
QY	2905	AGGCGACAGCAATTAATGTCGTATTAATGTCAGCAGCAGCAACAATTTCTCAGGGCTTTTG	2964
Db	7876	AGGCGACAGCAATTAATGTCGTATTAATGTCAGCAGCAGCAACAATTTCTCAGGGCTTTTG	7935
QY	2965	AGGCGCAACAGCATCTGTTCGCACTCAAGTCTGTGGGGCATCAAAACAGCTCCAGGCAAGAA	3024
Db	7936	AGGCGCAACAGCATCTGTTCGCACTCAAGTCTGTGGGGCATCAAAACAGCTCCAGGCAAGAA	7995
QY	3025	TCCTTGCTGTGGAAAGATACCTAAAGGATCAACAGCTCTCTGGGGATTTGGGGTTTGCCTCG	3084
Db	7996	TCCTTGCTGTGGAAAGATACCTAAAGGATCAACAGCTCTCTGGGGATTTGGGGTTTGCCTCG	8055
QY	3085	GAAAGACATTTTGGACACAGCTGCTGGCTTGGGAATGCTACTGTGGAGTAATAAATATCTTGG	3144
Db	8056	GAAAGACATTTTGGACACAGCTGCTGGCTTGGGAATGCTACTGTGGAGTAATAAATATCTTGG	8115
QY	3145	AACGAGTTTGGAAATTAACATGACCTCGATGGAGTGGAGGACAGAGAAATTAACAAATTTACACAA	3204

Db 8116 AACGATTTGGAAATACATGACCTGGATGGATGGACAGCAAAATTTACATTTACACAA 8175  
 QY 3205 GCTTAATACACTCTCTTAATTGAGAAATCGCAAAACGACGAAAGAAATGACAGAAAT 3264  
 Db 8176 GCTTAATACACTCTCTTAATTGAGAAATCGCAAAACGACGAAAGAAATGACAGAAAT 8235  
 QY 3265 TATTGGAATTAGATTAATGGCAAGTTTGGAATTGGTTTACATAACAATTTGGCTGT 3324  
 Db 8236 TATTGGAATTAGATTAATGGCAAGTTTGGAATTGGTTTACATAACAATTTGGCTGT 8295  
 QY 3325 GGTATATAAAATTTATATATAGATAGAGAGGCTGGTATGATTTAGAAATAGTTTGTG 3384  
 Db 8296 GGTATATAAAATTTATATATAGATAGAGAGGCTGGTATGATTTAGAAATAGTTTGTG 8355  
 QY 3385 CTGTACTTTCTATAGTAATAGATTAGGAGGATATTCACATTAATCTGTTTCAGACCC 3444  
 Db 8356 CTGTACTTTCTGATGTAATAGATTAGGAGGATATTCACATTAATCTGTTTCAGACCC 8415  
 QY 3445 ACCTCCCAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAAGAAAGAGTGGAGAGA 3504  
 Db 8416 ACCTCCCAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAAGAAAGAGTGGAGAGA 8475  
 QY 3505 GAGACAGACAGATCCATTGATTTAGTAAGACGATCTTACACTTATCTGGAGCAGATC 3564  
 Db 8476 GAGACAGACAGATCCATTGATTTAGTAAGACGATCTTACACTTATTTGGAGCAGATC 8535  
 QY 3565 TCGGAGACCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTGTTGATTAAACA 3624  
 Db 8536 TCGGAGACCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTGTTGATTGACGA 8595  
 QY 3625 GGATTGTGGAACCTCTGGAGCGGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCC 3684  
 Db 8596 GGATTGTGGAACCTCTGGAGCGGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCC 8655  
 QY 3685 TACAGTATGTGAGTACAGAACTTAAGAATAGTGTCTTAACTTCTCAATGCCACAGCCA 3744  
 Db 8656 TACAGTATGTGAGTACAGAACTTAAGAATAGTGTCTTAACTTCTCAATGCCACAGCTA 8715  
 QY 3745 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGATTAACAAGACCTTATAGACTA 3804  
 Db 8716 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGATTAACAAGACCTTATAGACTA 8775  
 QY 3805 TTGCGCACATACCTAGAGAAATAGACAGGCTTGAAGAGATTTCTCTAATAAGATGGT 3864  
 Db 8776 TTGCGCACATACCTAGAGAAATAGACAGGCTTGAAGAGATTTCTCTAATAAGATGGT 8835  
 QY 3865 GGCAGAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGGAAAGATGAGACGA 3924  
 Db 8836 GGCAGAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGGAAAGATGAGACGA 8895  
 QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGATCTAGA 3971  
 Db 8896 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCAGTATCTCGAGCCTAGA 8942

RESULT\_13  
 HIVPV22 9770 bp DNA linear VRL 01-OCT-1999  
 LOCUS Human immunodeficiency virus type 1, isolate PV22, complete genome  
 DEFINITION (H9/HTLV-III proviral DNA).  
 ACCESSION K02083  
 VERSION K02083.1 GI:555008  
 KEYWORDS TRK protein; acquired immune deficiency syndrome; complete genome; env protein; gag protein; long terminal repeat; pol protein; polyprotein; proviral gene; rev protein; reverse transcriptase; tat protein; transactivator.  
 SOURCE Human immunodeficiency virus 1 (HIV-1)  
 ORGANISM Human immunodeficiency virus 1  
 Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate lentivirus group.  
 REFERENCE 1 (bases 1 to 9770)  
 AUTHORS Muesling,M.A., Smith,D.H., Cabradilla,C.D., Benton,C.V., Lasky,L.A.

TITLE and Capon,D.J.  
 JOURNAL Nucleic acid structure and expression of the human  
 MEDLINE AIDS/Lymphadenopathy retrovirus  
 PUBMED Nature 313 (6002), 450-458 (1985)  
 85111157  
 2982104  
 REFERENCE 2 (sites)  
 AUTHORS van Beveren,C.P., Coffin,J. and Hughes,S.  
 TITLE Appendix B: HTLV-3/LAV genome  
 JOURNAL (in) Weiss,R.L., Teich,N., Varmus,H. and Coffin,J. (Eds.);  
 RNA TUMOR VIRUSES, MOLECULAR BIOLOGY OF TUMOR VIRUSES, SECOND  
 EDITION, 2: Supplements and Appendixes: 1106-1123;  
 Cold Spring Harbor Laboratory, CSH, NY (1985)  
 3 (bases 2111 to 2112)  
 REFERENCE Muesling,M.A.  
 AUTHORS unpublished  
 JOURNAL On Oct 4, 1994 this sequence version replaced gi:328551.  
 COMMENT [1] revised sequence, personal communication.  
 [(in) Weiss,R., Teich,N., Varmus,H. and Coffin,J. (Eds.);RNA Tumor  
 Viruses,Molecu review; bases 1 to 9769.  
 [3] revises [1],[(in) Weiss,R., Teich,N., Varmus,H. and Coffin,J.  
 (Eds.);RNA Tumor Viruses,Molecu].  
 This sequence for a H9/HTLV-III virus was determined from one  
 complete proviral clone [1]. Additionally, several cDNA clones of  
 the viral RNA were sequenced for comparison with the entire  
 proviral sequence. The differences between cDNA and proviral DNA  
 are extensive and are listed in the Sites table as variations. The  
 authors believe that the variations may be due in part to different  
 strains in the H9/HTLV-III cell line, because it was established by  
 infection with material from several AIDS patients. With the  
 addition of g at 2111, gag cds and pol cds are very close to those  
 of HXB2, BRU, and related HIV viruses.  
 For details and other references pertaining to Sites and Features,  
 see the HIV reference entry.  
 location/Qualifiers  
 1..9770  
 /organism="Human immunodeficiency virus 1"  
 /proviral  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:11676"  
 10..643  
 /note="5' LTR"  
 386..395  
 /bound\_molecly="Sp1 III"  
 397..406  
 /bound\_molecly="Sp1 II"  
 408..417  
 /bound\_molecly="Sp1 I"  
 463..560  
 /note="R repeat 5' copy"  
 464..9678  
 /note="genomic mRNA"  
 464..9678  
 /note="tat, rev, nef subgenomic mRNA"  
 464  
 /note="numbered 1 in [1]"  
 510  
 /note="a in provirus; g in cDNA [1]"  
 /replace="g"  
 575  
 /note="g in provirus; a in cDNA [1]"  
 /replace="a"  
 645..662  
 /bound\_molecly="Lys-rRNA primer"  
 753..5822  
 /note="tat, rev, nef subgenomic mRNA intron 1"  
 799..2337  
 /note="gag polyprotein precursor"  
 /codon\_start=1  
 /protein\_id="AAB59866.1"  
 /db\_xref="GI:328554"  
 /translation="MGARASVYSGGELDRMEKIRLRPGKKKKYKIKHIVASRELEPR  
 AVNPGLETSBGCRIILGOLQPSLQTSSELRSLRYNTVATLYCYHRIEIKDKREALD

KITEEONKSKKKAQAAADGTHSSVQSNVYIVONIGQWHAQISPRITLNAVXYVE  
 EKAPEASPEVTPMHSALSGEATPODLTNTLNTVGHQAOAMQMLKETINERAAEMDWBHVH  
 HAGPIAPSOEERREPSDIAGTSTLQEDIGMNTNPPIDVEIYKRMILGLNLTAVP  
 YSPSTILDIROGKEPFRPDYLDREFYKTLIRADAQSEVNMNTEFLVONAPDCITLIL  
 KALGAATLIEEMMTACQGVGGHARVLAFLAASQVITNATIMQGRNORNRQKAGC  
 FNGCEGHARNCRAPKRKCGMKCGKECHKDCTERANFLKGLIMPYSYKGRCPENFLO  
 SPPETPAPELQSRPEPTAPPEBSFRSGVEYTTTPQOKEPIDKELYPULTSLRSLFGND  
 PESSQ"  
 CDS  
 2094. .5141  
 /partial  
 /note="pol polyprotein (NH2-terminus uncertain)"  
 /codon\_start=-1  
 /protein\_id="AAB59867.1"  
 /db\_xref="GI:328555"  
 /translation="PFEDLAFLOGKAREFSSEQTRANSPITISSEQTRANSPTRREIQ  
 VNGVRNNSSEAGADROGTSVSNFPQITLMORPLVTIKIGGOLKEALDTCADQTVLE  
 ESSLGRMSPKMGIGIGGIFKRYOYDOLLIEIGCKAIGTVLGPPTNITGRLLTQO  
 IGCTIANPISPLEIYPVYKLRQMDGPYKQMPLEBEKIKALVELICTEMEKGKISKIGI  
 PENPVPFALFKKDDSTYKRWKLVDFEYRLNKTQDFWVVOGIIHPGALKRKRKYVL  
 DNGDAVFSPVDEDEDFRYKTAFTIPISINNETGIRQYVNVLPQMGKSPALFQSSMTKIL  
 LEPEKQNDIVIQYMDLIVYSDLEIGHRTKIEELRQHLNMGLTTPPKRKQKPEP  
 PLPMGVEHLDPKMTVQPIVLEPKDSMTVNDIQKLYGLNMAISOYIGIKYKQCKLL  
 RQTKLTFEPTLEAELEIAENREILKEPVHAYGDSKDLIAEIOKQOQCMQTYDIO  
 YOEPRKNLKTGKARKRGARHTNDKQLTEAVQKITTESIVTWGTPKRLDTPQKTEWE  
 TWMTWQATWIPMEVEYNTPPLVYKLYMTQLEKEPIVGAETYYVDGANRETRIGKAGY  
 LYNKROKRVPLTNTNQKTELQAIYLAQLDSGLEVNITVDSOYALGIIQAOPQSESKIT  
 ELNRQNIIEELIKOKVYLAWPANHGKIGISNOYDKLVASGIRKILFDGIDKADDEHSE  
 KXHSMMARASADFNLPVYAKEIVASQCSQCKQKGAEMGVQDQSPGIMQDCPLDEKKA  
 VLVAVHAASQVIEAEVPIPAETGOETFAVELIKLGRMPVYKTIHNDGNSNPTSAVYKA  
 CMMACITKQEFQIPYNPQSQGVESNNKELKKTICGVYRQAEHLTAVQMAVFINFRK  
 KGGIGYSNGEKIYDIIATDITQKLELQKTIQINFYRYRDSNSPLMKGPAKLMLMG  
 ECAVYIQDSDIKYVRRKAKIIRDYGRQMGADDCVASRODED"  
 2111. .2112  
 /citation=[1]  
 5086. .5664  
 /note="vif protein"  
 /codon\_start=-1  
 /protein\_id="AAB59868.1"  
 /db\_xref="GI:328556"  
 /translation="MENRMQVAVIWMQVDRMRILRTMKSLVKHMHYVSGKARGMEYRHHY  
 ESRPIRSEVHIVPIGLDARIVITVYWGILHGERPMHICQGSYEMRRKRYSTQVDEPL  
 ADQILHLYFDQCFSDSARIKALIGHIVSPREIYQAGNNKVSLOYLALALITPKKIK  
 PLPVSYTKLEDRNMKPKTKYGRHSHTMNH"  
 5604. .5840  
 /note="vpr protein"  
 /codon\_start=-1  
 /protein\_id="AAB59869.1"  
 /db\_xref="GI:328557"  
 /translation="WQAPEDQGPQREPHNEMTLELELNKNEAVRHPRITWHLGLQO  
 HLYETVGTWAGVEAIIIRLQDLFIHQNVWST"  
 5716  
 /note="g in provirus; a in cDNA [1]"  
 /replace="a"  
 join(5876. .6090,8421. .8466)  
 /note="tat protein"  
 /codon\_start=-1  
 /protein\_id="AAB59870.1"  
 /db\_xref="GI:328552"  
 /translation="MEPYDPLREPMKHGSSQPKTACTNCTCYKCCCFHCQVCFITKALG  
 ISTYGRKKRRQRRRPPQGSQTHOVSLSKOPTSQKPDTPGKE"  
 5992  
 /note="a in provirus; g in cDNA [1]"  
 /replace="g"  
 6007  
 /note="c in provirus; t in cDNA [1]"  
 /replace="t"  
 join(6015. .6090,8421. .8695)  
 /note="rev protein"  
 /codon\_start=-1  
 /protein\_id="AAB59871.1"  
 /db\_xref="GI:328553"  
 /translation="WAGRSSGSDSDLLKAVRLIFELYQSNPNPPEGTROARRNNRRR  
 WEROROIHSIERILSTYTGKSAEPVLDLPLEPRLTILDCNECCGSGIGQGVSSPDI

Query Match	Best Local Similarity	Matches 2840:	Conservative	0:	Mismatches	41:	Indels	6:	Gaps	1:																																		
variation	6047	/note="c in provirus; g in cdna [1]"	/replace="g"	6051	/note="c in provirus; a in cdna [1]"	/replace="a"	6055	.6057	/note="agg in provirus; gaa in cdna [1]"	/replace="gaa"	6091	.8420	/note="tat cds intron 2"	6091	.8420	/note="rev cds intron 2"	6091	.8420	/note="tat, rev, nef subgenomic mRNA intron 2"	6107	.6352	/note="vpu protein"	/codon_start=1	/protein_id="AAB59872.1"	/db_xref="GI:328558"	/translation="MOPQIAIVALLVAIIIAIILWSIVLIEYKKILRQKIDRLIDR LIERAEDSGNESEGEISALVEMGVEMHRAHPMDVDDL"	6108	/note="t in provirus; c in cdna [1]"	/replace="c"	6120	/note="a in provirus; c in cdna [1]"	/replace="c"	6125	.6126	/note="gc in provirus; gtaac in cdna [1]"	/replace="gtaac"	6136	/note="a in provirus; c in cdna [1]"	/replace="c"	6235	/note="t in provirus; a in cdna [1]"	/replace="a"	6267	.8837
variation	6265	CAATGAGGTAAAGAGAAATATATACGACTTGTGAGATGGGGTGAAGATGGGGACCA	6324	1351	TGCTCCTTGGGATATTGATGATCTGTAGTGTCTACAGAAAAATTGTGGTCCACCGCTATT	1410	6325	TGCTCCTTGGGATGTGATGATCTGTAGTGTCTACAGAAAAATTGTGGTCCACCGCTATT	6384	1411	ATGGGGTACCTGTGTGAGAGAACACACACCTCTATTTTGTGCATCAGATGCTTAAG	1470	6385	ATGGGGTACCTGTGTGAGAGAACACACACCTCTATTTTGTGCATCAGATGCTTAAG	6444	1471	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCCA	6504	6445	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCCA	6504	1531	ACCCACAAGAGTATGTTGTTAAATGTGACAGAAAAATTTTAAACATGTGAAAAATGACA	1590	6505	ACCCACAAGAGTATGTTGTTAAATGTGACAGAAAAATTTTAAACATGTGAAAAATGACA	6564	1591	TGTTAGAACAGATCAGATGAGATATTAATCACTTTATGGGATCAAAAGCCCTAAAGCATGTG	1650	6565	TGTTAGAACAGATCAGATGAGATATTAATCACTTTATGGGATCAAAAGCCCTAAAGCATGTG	6624	1651	TAAATTAACCCCACTGCTGTGTTAGTTTAAAGTCCACGATTTGAAGATGATCTATA	1710	6625	TAAATTAACCCCACTGCTGTGTTAGTTTAAAGTCCACGATTTGAAGATGATCTATA	6684	1711	CCATTAAGTACAGAGGAGATGATTAATGAGAAAGAGAGATTAATAAAGCTGCTTTCA	1770		





RESULT 14  
LOCUS E01088 3156 bp DNA linear PAT 29-SEP-1997  
DEFINITION Nucleic acid sequence coding for HTLV envelop protein.  
ACCESSION E01088  
VERSION E01088.1 GI:2169347  
KEYWORDS JP 1987012799-A/1.  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE  
AUTHORS 1 (bases 1 to 3156)  
Robaoto,M.K., Robaato,C.G., Eragamu,P.R., Jiyooji,M.S. and  
Furotsushii,I.U.  
TITLE ENVELOPE PROTEIN OF VIRUS OF ACQUIRED IMMUNODEFICIENCY SYNDROME  
JOURNAL Patent: JP 1987012799-A 1 21-JAN-1987;  
F. HOFMANN LA ROCHE & CO AG, USA GOVERNMENT  
COMMENT OS Human [Homo sapiens]  
PN JP 1987012799-A/1  
PD 21-JAN-1987  
PF 18-APR-1986 JP 1986089830  
PI 19-APR-1985 US 85 725021  
PI ROBATO MITSUCHIERU KUROURU, ROBATO CHIYARUZU GARO, PI  
ERAGAMU PUREMUKUMAA REDEL, JIYOOJI METSUDO SHIYUO, PI FUROTSUSHII  
IICHINGU UONGUSUTARU  
PC C07K13/00,A61K39/21,C07H21/04,C12N1/20,C12N15/00,C12P21/00,PC  
C12P21/02,  
PC G01N33/569,G01N33/577,(C12N1/20,C12R1:125),(C12N1/20,C12R1:19),PC  
(C12P21/02,  
PC C12R1:125),(C12P21/02,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: library=genome library;  
FH key Location/Qualifiers  
FT mat\_peptide 1..3157  
FT Location/Qualifiers  
1..3156  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 1095 a 536 c 764 g 761 t  
ORIGIN

Query Match 41.7%; Score 2598.2; DB 6; Length 3156;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2637; Conservative 0; Mismatches 38; Indels 6; Gaps 1;

QY 1291 CCATGAGAGTGAAGAGAGATATCAGCACTTGTGGAGATGGGGGTGGAATGGGGACCA 1350  
D 476 CAATGAGAGTGAAGAGAGAAATATCAGCACTTGTGGAATGGGGGTGGAATGGGGACCA 535  
QY 1351 TGCCTCTTGGGATATTATGATCTGTAGTGTACAGAAAATTTGGGTCACTGTAATT 1410  
D 536 TGCCTCTTGGGATCTTATGATCTGTAGTGTACAGAAAATTTGGGTCACTGTAATT 595  
QY 1411 ATGGGAGTACCTGTGTGAAGAGCAACCACTATTTTGGCATCGATCGCTTAAG 1470  
D 596 ATGGGAGTACCTGTGTGAAGAGCAACCACTATTTTGGCATCGATCGCTTAAG 655  
QY 1471 CATATGATACAGAGTACATATGTTTGGGCACACATGCGTGTACCCACAGACCCA 1530  
D 656 CATATGATACAGAGTACATATGTTTGGGCACACATGCGTGTACCCACAGACCCA 715

QY 1531 ACCCACAAGAAGTAGTATTGTGTAATGTGACAGAAAATTTTAACATGTGAAAAATGACA 1590  
D 716 ACCCACAAGAAGTAGTATTGTGTAATGTGACAGAAAATTTTAACATGTGAAAAATGACA 775  
QY 1591 TGGTAGAACAGATGATGAGATATTAATCACTTTAGGATTCAAAGCCTTAAGCCATGTG 1650  
D 776 TGGTAGAACAGATGATGAGATATTAATCACTTTAGGATTCAAAGCCTTAAGCCATGTG 835  
QY 1651 TAAATTAACCCCACTGTGTAGTTAATTAAGTGCACCTGTGTTGAACAAATGATACATATA 1710  
D 836 TAAATTAACCCCACTGTGTAGTTAATTAAGTGCACCTGTGTTGAACAAATGATACATATA 895  
QY 1711 CCAATAGTAGAGGGGAGAGATGATAATGAGAAAGAGAGATAAAAAAGCTGCTTTCA 1770  
D 896 CCAATAGTAGAGGGGAGAGATGATAATGAGAAAGAGAGATAAAAAAGCTGCTTTCA 955  
QY 1771 ATATCAGACACAGATTAAGAGATTAAGTGCAGAAAGAAATATGCACTTTTATTAACCTTG 1830  
D 956 ATATCAGACACAGATTAAGAGATTAAGTGCAGAAAGAAATATGCACTTTTATTAACCTTG 1015  
QY 1831 ATATAGTACCAATAGATAA-----TACAGCTATAGTTGATTAAGTGAACACCTCAG 1884  
D 1016 ATATATACCAATAGATTAAGATTAAGTACCTACGCTATACGTTGACAAAGTTGACACCTCAG 1075  
QY 1885 TCATTAACACAGGCGCTGCCAAGGATCCTTTGAGCCCAATTCATATTAATGTCGCC 1944  
D 1076 TCATTAACACAGGCGCTGCCAAGGATCCTTTGAGCCCAATTCATATTAATGTCGCC 1135  
QY 1945 CGGCTGTTTGGGATTTCTAAATGTAAATTAAGAGCTTCATGGAACGAGACCATGTA 2004  
D 1136 CGGCTGTTTGGGATTTCTAAATGTAAATTAAGAGCTTCATGGAACGAGACCATGTA 1195  
QY 2005 CAAATGTACACAGTACATATGTAACATGGAATCAAGCCAGTATGTAACATCACTGC 2064  
D 1196 CAAATGTACACAGTACATATGTAACATGGAATCAAGCCAGTATGTAACATCACTGC 1255  
QY 2065 TGTTAATGCGAGTCTAGCAGAAAGAGATGTAATTAAGTATGATCCCAATTAACAGACA 2124  
D 1256 TGTTAATGCGAGTCTAGCAGAAAGAGATGTAATTAAGTATGATCCCAATTAACAGACA 1315  
QY 2125 ATGCTAAACCAATATGTAACGTGACGTCGACACATCTGTGAAGAAATTAATGTACAAAGCCA 2184  
D 1316 ATGCTAAACCAATATGTAACGTGACGTCGACACATCTGTGAAGAAATTAATGTACAAAGCCA 1375  
QY 2185 ACAACATTAACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGAGACATTTGTTACAA 2244  
D 1376 ACAACATTAACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGAGACATTTGTTACAA 1435  
QY 2245 TAGGAAAAATAGGAATATGAGACAGACATTTGTAATTAAGTATGAGCAAAATGGAATG 2304  
D 1436 TAGGAAAAATAGGAATATGAGACAGACATTTGTAATTAAGTATGAGCAAAATGGAATG 1495  
QY 2305 CCACTTAACAAAGATAGTACCAATTAAGAGAACATTTGCAAAATTAACAAATTA 2364  
D 1496 CCACTTAACAAAGATAGTACCAATTAAGAGAACATTTGCAAAATTAACAAATTA 1555  
QY 2365 TCTTTAAGCAATCCTCAGAGAGGAGCCAGAAATTTGTAACGACAGTTTAATTTGAGAG 2424  
D 1556 TCTTTAAGCAATCCTCAGAGAGGAGCCAGAAATTTGTAACGACAGTTTAATTTGAGAG 1615  
QY 2425 GGGATTTTCTACTGTAATTCACACACACTGTTTAATTAAGTATGATGTTTAAATAGTACTT 2484  
D 1616 GGGATTTTCTACTGTAATTCACACACACTGTTTAATTAAGTATGATGTTTAAATAGTACTT 1675  
QY 2485 GGAATAGTGAAGGGTCAATTAACACTGAAGAGTGCACAATCAGCTCCATGCGAA 2544  
D 1676 GGAATAGTGAAGGGTCAATTAACACTGAAGAGTGCACAATCAGCTCCATGCGAA 1735  
QY 2545 TAAACCAATTTAATTAACATGTGCAGAGAAATAGGAAAGCAATGATGCCCCCTCCATCA 2604  
D 1736 TAAACCAATTTAATTAACATGTGCAGAGAAATAGGAAAGCAATGATGCCCCCTCCATCA 1795  
QY 2605 GTGACAAATTAAGTATGATGTTCAATTAATTAATTAATGAGTATTAACAAAGAGATGTTGTA 2664

Db	1796	GCAGCAAAATAGATGTTTCATCAATATATACAGGGCTGCTATTAAACAAGAGATGGTGTGA	1855
Qy	2665	ATAACAACATGGGTCCGAGATCTTCAGACCTGGAGAGCGGATATAGAGCAATTGGGA	2724
Db	1856	ATAACAACATGGGTCCGAGATCTTCAGACCTGGAGAGCGGATATAGAGCAATTGGGA	1915
Qy	2725	GAAGTGAATATATAAATATATAAAGTAAATAATGAAACATTGAGAGTACCCACCA	2784
Db	1916	GAAGTGAATATATAAATATATAAAGTAAATAATGAAACATTGAGAGTACCCACCA	1975
Qy	2785	AGGCAAGAGAGAGAGTGGTCCAGAGAGAAAAAGAGAGTGGGAATAGAGCTTTGTCC	2844
Db	1976	AGGCAAGAGAGAGTGGTCCAGAGAGAAAAAGAGAGTGGGAATAGAGCTTTGTCC	2035
Qy	2845	TTGGGTTCTTGGAGCAGCAGAGGAAGCACTATGGCTCAGCTCANTAGCGTGAAGTAC	2904
Db	2036	TTGGGTTCTTGGAGCAGCAGAGGAAGCACTATGGCTCAGCTCANTAGCGTGAAGTAC	2095
Qy	2905	AGGCGACAGCAATTTGCTGATATAGTGAAGCAGCAAGCAATTTGCTGAGGGCTATG	2964
Db	2096	AGGCGACAGCAATTTGCTGATATAGTGAAGCAGCAAGCAATTTGCTGAGGGCTATG	2155
Qy	2965	AGGCGACAGCAATTTGCTGATATAGTGAAGCAGCAAGCAATTTGCTGAGGGCTATG	3024
Db	2156	AGGCGACAGCAATTTGCTGATATAGTGAAGCAGCAAGCAATTTGCTGAGGGCTATG	2215
Qy	3025	TCCTGGCTGTGGAAAGATACCTTAAGATCAACAGCTCTGGGGATTTGGGGTCTCTG	3084
Db	2216	TCCTGGCTGTGGAAAGATACCTTAAGATCAACAGCTCTGGGGATTTGGGGTCTCTG	2275
Qy	3085	GAAGACATTTGGACACCTGCTGCTGATGATGCTAGTGGAGTAAATTAATCTCTGG	3144
Db	2276	GAAGACATTTGGACACCTGCTGCTGATGATGCTAGTGGAGTAAATTAATCTCTGG	2335
Qy	3145	AACAGATTTGGAATATACATGACCTGATGAGTGGAGCAGAGAAATTAACATTAACAA	3204
Db	2336	AACAGATTTGGAATATACATGACCTGATGAGTGGAGCAGAGAAATTAACATTAACAA	2395
Qy	3205	GCTTAATACACTCCTTAATGAGAAATGCGCAAAACAGCAAGAAAGATGACAGAAAT	3264
Db	2396	GCTTAATACACTCCTTAATGAGAAATGCGCAAAACAGCAAGAAAGATGACAGAAAT	2455
Qy	3265	TATTTGAATAGATTAATGGGCAAGTTTGGAGTTTGGTTAAATTAACAAATGGCTGT	3324
Db	2456	TATTTGAATAGATTAATGGGCAAGTTTGGAGTTTGGTTAAATTAACAAATGGCTGT	2515
Qy	3325	GATTAATAAATTTATCATATAGTATAGAGAGGCTTGGTGGTTTAAAGATAGTTTGG	3384
Db	2516	GATTAATAAATTTATCATATAGTATAGAGAGGCTTGGTGGTTTAAAGATAGTTTGG	2575
Qy	3385	CTGACTTTCTATAGTAATAGATTAGGAGGGATATCCACTTATCTGTTAGAGCC	3444
Db	2576	CTGACTTTCTATAGTAATAGATTAGGAGGGATATCCACTTATCTGTTAGAGCC	2635
Qy	3445	ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCGAAGAAATGAAGAAAGAGTGAAGGA	3504
Db	2636	ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCGAAGAAATGAAGAAAGAGTGAAGGA	2695
Qy	3505	GAGACAGAGACAGATTCATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	3564
Db	2696	GAGACAGAGACAGATTCATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	2755
Qy	3565	TGCGGAGCCTGTGCTTTCAGTACACCGCTTGAGAGACTTACTCTTGAATTTAAAGA	3624
Db	2756	TGCGGAGCCTGTGCTTTCAGTACACCGCTTGAGAGACTTACTCTTGAATTTAAAGA	2815
Qy	3625	GGATTTGGAACCTTCGAGCGAGCGAGGGGTGGGAAGCCCTTAATATTTGGGAAATCC	3684
Db	2816	GGATTTGGAACCTTCGAGCGAGCGAGGGGTGGGAAGCCCTTAATATTTGGGAAATCC	2875
Qy	3685	TACAGTATTTGAGTACAGAACTAAGAAATAGTCTGTTAACTTCTCAATCCACAGCA	3744
Db	2876	TACAGTATTTGAGTACAGAACTAAGAAATAGTCTGTTAACTTCTCAATCCACAGCA	2935
Qy	3745	TAGCAGTACCTGAGGAGCAGCATAGGCTTATAGAAAGTATTAACAGCAGCTTATGACCTA	3804
Db	2936	TAGCAGTACCTGAGGAGCAGCATAGGCTTATAGAAAGTATTAACAGCAGCTTATGACCTA	2995
Qy	3805	TTGCGCACATATACATAGAAAGTAAAGACAGGCTTGGAAGAGTATTTCTTAAATGGGT	3864
Db	2996	TTGCGCACATATACATAGAAAGTAAAGACAGGCTTGGAAGAGTATTTCTTAAATGGGT	3055
Qy	3865	GGCAAGTGTCTCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG	3924
Db	3056	GGCAAGTGTCTCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG	3115
Qy	3925	GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTGAGA	3965
Db	3116	GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCAGTATCTGAGA	3156
RESULT 15			
HIVBH102			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SEGMENT			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
Original source text: Human Immunodeficiency virus type 1 (HIV-1)			
proliferative DNA clone BH10.			
[(in) Weiss, R., Teich, N., Varmus, H. and Coffin, J. (Eds.), RNA Tumor			
Viruses, Molec. review, bases 1 to 8932.			
The BH10 sequence differs from BH8 and BH5 by 0.9% in the coding			
regions and 1.8% in the noncoding regions, and the authors of [1]			
believe that these are stable variants.			
The HTLV-III genome encodes at least seven proteins: gag, pol, env,			
tat, trs, 27K antigen and the src 23K product. The 3' ORF			
(positions 8153-8773) is truncated in BH10 (stop codon at positions			
8522-8524), but reads through in BH8 and other sequences to yield			
what is now called the 27K antigen.			
The mechanism for pol gene translation has not been elucidated: a			
gag-pol fusion protein is possible; splicing or frameshift have not			
been ruled out. The viral protease would be determined by the			

region in question.  
 The Tat protein (trans-activator protein, approximately 14 kb) is an effector of an autostimulatory pathway through interaction with a positive control element, the trans-activating responsive sequence, TAR. Tat seems to be a transcriptional control molecule in HTLV-1, but is both that and a post-transcriptional regulatory molecule in HTLV-III. Deletion mutants in the tat gene are incapable of prolific replication and exhibit no cytopathic effects in T4 cell lines.  
 In addition to the  
 9.4 kb genomic mRNA, subgenomic mRNAs of 7.4, 5.5, 5.0, 4.3, 2.0 and 1.8 kb have been detected.

## FEATURES

## source

Location/Qualifiers  
 1. .8932

/organism="Human immunodeficiency virus 1"

/mol\_type="genomic RNA"

/db\_xref="taxon:11676"

prim\_transcript <1..>8932

/note="genomic mRNA"

prim\_transcript <1..>8932

/note="tat, rev, nef subgenomic mRNA"

## intron

/note="tat, rev, nef subgenomic mRNA intron 1"

## CDS

/note="gag polypolypeptide precursor"

/codon\_start=1

/protein\_id="AA44201.1"

/db\_xref="GI:326388"

/translation="MGARASVLSGGLDREKRIIRLRGGKKRYKIKHITWASRELEF  
 AVNPGLTSEGCROILGLOPSELTSSELRLYNVALVCVHORIEIKDKLEAD  
 KLEEDNKRKKAQAQAADGSHSOVSQNPVIONOGVHQAISPTANWAVVE  
 EKAFSEIPVIMFALSAGATODLNTMLNTVNGHOAMOMAKETINDEAEWVRVHV  
 HAGPIAFGQREPRSDIAGTSTLQOIGMNTMNPVIGSEIYKRWIIIGLANKIVM  
 YSPSTLIDIRGKRPDRVDVDFYKTLRAEOASOEKMMTELLVONNPPCKTL  
 KAIGPATLEEMMTACOGVGGPRHKAIVLAEMASQVNTNTIMORGNFNRKMKVC  
 PNCGKSHTRKRCRAPKRCGKMGKGHOMKQTEQANFLGKIMPSYGRPNLQ  
 SREPTAPPLQSPREPTAPPEESFRSGVETPTTPQKQEPIDKELPLTSLRSLFND  
 PSSQ"

1407..4454

/gene="pol"

<1407..4454

/note="ORF for pol polypeptide containing stop codons"

/codon\_start=1

/product="pol polypeptide"

/protein\_id="AA44198.1"

/db\_xref="GI:326385"

/translation="FRREDLAFIQAQKARESSQETRANSPPTISSEQTRANSPTRREIQ  
 VWRDNNSPEAGADROGVSNFNPQITLMORPLVITIKIGQLKEALLDGDADTVLE  
 EMSLPGRMKPKMIGIGGIFKVRQYDQILIEIGHKAIGVLVGPPTVNTIGRLTLQ  
 IGTLPPIPIETVPPVYKLRPGMDGPKVOMPLTEERIKALVEICTEMEKEKISKIG  
 PENPYNTPVPAIKKDDSTKRWIKLDPRELKRPDDEWVLOGIPIHAPGLAKKKSATVL  
 DNGDAYSVPLDDEFRKYTAFTTIPSTINNEPGRFYQVNVPOCKGSPALFOSSMTKI  
 LEFFKRONPDIIVYQIMDDLVSQDEIGHRKIEELRHLRLMGLTTPDKHOKRP  
 PFLMGEELHPRDKWTQPIVLPKDSWTVNDIQKLVGLKMAASQIYIGIVROJCKLL  
 RGRKALTEVPLTEAELELAENREILKEPVHGYDPSKDLAELQKQOGQCTYQI  
 YQEPFKLKTGKYARMGANTNDVOKLTEAVOKTTESIVYMGKTPKFKPIKQKEME  
 TMMTEWQATWIMEWEFVNTPLVYKLMYOLEKEPIVGAETFEVYGAAANRETKISGAKY  
 VTNGKQKVPPLINTTNOKTELOAIYIALADSGLENAVITDSOYALCITQAOQPKSS  
 ELVNQILEQLEKEVYLAWPAHKGIGGNEQYDKLASIGIRILFLDGIQDKODEHE  
 KYHNSMRMAASDNLPRVVAKEIVASCDKQLGKGAHQGVDCSPGIMQDCKLDEBE  
 VILVAHVASGYIEAEVIPAETQETAYFLIKLAGRPVQTIHTDNGSNFTSIVKKA  
 CMMAGIKOEFGIPYNPOSGVESMKNELKKIIGOVDOAEHLKTAQVMAVFLHNEFR  
 KGGIGGSAGERIVDITADTQIKLOKQTKTIONPFVYVYRDSRNPJMKRPALLMKG  
 EGAIVIDNSDIDIVPRRRKAKITRDYKQAGADCAVSRQDED"

4399..4977

/note="vif protein"

/codon\_start=1

/protein\_id="AA44202.1"

/db\_xref="GI:326389"

/translation="MENKROVMYVMOVDNRKIRITKWSLVYKHMYVSGKANGMYRIHY  
 ESPHPRISSEVHIPLGDARLVITTYWGLTGERDWHLGQGVSIEMRRKRYSTVDPEPL  
 ADQILHLYPDCFSDSAIRKALLGHIVSPRCEYOAGHNKVGSLQYIALAALITPKKIK

## CDS

PLPSTYKLTEDENMRPKQTKGHSHTNGH"  
 4917..5153

/note="vpr protein"

/codon\_start=1

/protein\_id="AA44203.1"

/db\_xref="GI:326390"

/translation="MEQAPEDQGPQREPHNWTLELLEIKNEAVRHFPRIWHLGIGQ  
 HYEYGDPMAGVEAIRILOQLFIHFQWVST"

join(5189..5403,7734..7779)

/note="tat protein"

/codon\_start=1

/protein\_id="AA44199.1"

/db\_xref="GI:326386"

/translation="MEPVRLPEPMKHPGSOPTACTNCGYCKKCFHCQVFTKALG  
 ISYGRKKRORRRRPPGOSQTHVSLSKQPSQSGRDPGPKK"

<5189..5403

/note="tat protein, (first expressed exon)"

/number=2

join(5328..5403,7734..8008)

/note="rev protein"

/codon\_start=1

/protein\_id="AA44200.1"

/db\_xref="GI:326387"

/translation="MAGRSQSDSEDLKAVRLIKELYQSNPPNPPECTQARRRRR  
 WRERQRIHSISERILSTYLGRSABEVPLQPLERLTLDNEDCGTSGTGVGSPQI  
 LVSEPVLESQAKE"

<5328..5403

/note="rev protein, (first expressed exon)"

/number=2

5404..7733

/note="tat cds intron 2"

5404..7733

/note="rev cds intron 2"

5404..7733

/note="tat, rev, nef subgenomic mRNA intron 2"

5420..5665

/note="vpu protein"

/codon\_start=1

/protein\_id="AA44204.1"

/db\_xref="GI:326391"

/translation="MQPIQAIYALVAVYAIITAVWSIYIIEYRKILRQKIDRLDR  
 LIEAEDSGNESEGETSALVEKGVEMHNPWVDL"

5580..8150

/note="envelope polypeptide"

/codon\_start=1

/protein\_id="AA44205.1"

/db\_xref="GI:326392"

/translation="MRVKEKYQHLMRGWRMGTMGLGMLICATEKLVNTYVYQVY  
 WKBATYTLFCASAKAYDTEVHNVMAFHACVPPDPNPQEVVLAIVTEENFMMKNDVVE  
 OMHEDIISLMDQSLKPCVKTLPLCVSLKCDLNDNTNSSGRLMEKEIKNCSEFN  
 ISHSIGKVOKEEYAFYKLDIIPIDNDTSTYTLSCNTSVITQACRVSPEPIHYC  
 APAGPQIKNNKNTFNGTCTCTVNVYQCHGIRBPVYQDLLNGSLAESEVYIRSAN  
 FTMNAKTIIVQLNDSVEINCTPRNNNTFRKIRIRORGBRFPYITGKIGNRRQAHNIS  
 RAWMNNTLKQIDSKLRQEGNNKTIIFKQSSGDPETVHSFNGGSEFFYCNSQTLN  
 STFNSTWSTKSGNNTEGSDTITLPLRIKQIIMMOEYGAAPISQIRGSSWIT  
 GLTLTNGKVSNNSEIFFRGGGDMNDMSSELYKKKVAIEELGVAAPKRRVVR  
 EKAVAGIGALFIEFLGAGSTWGAASWTIVQAROLLSGIVOOONMLRAIEQOHL  
 QLVWQKQLOALFILLAEVRLKQOOLGIGWCSGKLICTPAVPMNASWSKSIQELTN  
 NMTWMEHREINNTSILSHLIESQNOQKRNQELLEDKMASLNNWPIITWMTYI  
 KLPIMTVGGLVGRIVFAVILSVNRRQGSPLSFQTHLPITPGPDRPGIEEGGER  
 DDRORSIRLVNGSLALIMDDIRSLFSGHYRLDRLILVRIIVELGRGREALKYWN  
 LLQWSEQLKNSAVILLNATAIVAAETQDRIIVVOGAVRAIRHRIIRILOGIERILL  
 "

7734..>8008

/note="rev protein"

/number=3

7734..>7779

/note="tat protein"

/number=3

8152..8523

/note="nef protein, (first expressed exon; premature  
 termination)"

/number=3

## CDS

/note="vif protein"

/codon\_start=1

/protein\_id="AA44202.1"

/db\_xref="GI:326389"

/translation="MENKROVMYVMOVDNRKIRITKWSLVYKHMYVSGKANGMYRIHY  
 ESPHPRISSEVHIPLGDARLVITTYWGLTGERDWHLGQGVSIEMRRKRYSTVDPEPL  
 ADQILHLYPDCFSDSAIRKALLGHIVSPRCEYOAGHNKVGSLQYIALAALITPKKIK

## CDS

/note="vif protein"

## exon

## CDS

## CDS

```
/codon_start=1
/protein_id="AA44206.1"

Query Match      41.7%; Score 2597.8; DB 14; Length 8932;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2639; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

OY 1291 CCATGAGGTGAGAGAGAAATATGACGACCTGTGTGAGATGGGGGTGGAATGGGACCA 1350
    |||||||
Db 5578 CAATGAGGTGAGAGAGAAATATGACGACCTGTGTGAGATGGGGGTGGAATGGGACCA 5637

OY 1351 TGTCTCTGGGATATGTGATCTGTAGTGTACAGAAAATTTGGGTCAACGCTATT 1410
    |||||||
Db 5638 TGTCTCTGGGATGTGTATGATCTGTAGTGTACAGAAAATTTGGGTCAACGCTATT 5697

OY 1411 ATGGGATCTGTGTGAGAGAGCAACCACTCTATTTTGTGATCAGATGCTAAAG 1470
    |||||||
Db 5698 ATGGGATCTGTGTGAGAGAGCAACCACTCTATTTTGTGATCAGATGCTAAAG 5757

OY 1471 CATATGATACAGAGGTACATATATGTTGGGCGACACATGCTGTACCCAGACCCCA 1530
    |||||||
Db 5758 CATATGATACAGAGGTACATATGTTGGGCGACACATGCTGTACCCAGACCCCA 5817

OY 1531 ACCCAAGAGATAGTATGTGTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 1590
    |||||||
Db 5818 ACCCAAGAGATAGTATGTGTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 5877

OY 1591 TGGTAGAACATATGATGATATATCACTTTATGGGATCAAAAGCCTAAAGCCTATG 1650
    |||||||
Db 5878 TGGTAGAACATATGATGATATATCACTTTATGGGATCAAAAGCCTAAAGCCTATG 5937

OY 1651 TAAATTTAACCCACTCTGTGTAGTTAAAGTGCACTGATTTGAAGATGTACTAATA 1710
    |||||||
Db 5938 TAAATTTAACCCACTCTGTGTAGTTAAAGTGCACTGATTTGAAGATGTACTAATA 5997

OY 1711 CCAATAGTAGAGCGGGAGAAATGATTAATGAGAAAAGAGATAAAAACTGCTTTTCA 1770
    |||||||
Db 5998 CCAATAGTAGAGCGGGAGAAATGATTAATGAGAAAAGAGATAAAAACTGCTTTTCA 6057

OY 1771 ATATCAGACAAGCATTAAGATTAAGTGTGCAAGAAATATGCAATCTTTTATAACTTG 1830
    |||||||
Db 6058 ATATCAGACAAGCATTAAGATTAAGTGTGCAAGAAATATGCAATCTTTTATAACTTG 6117

OY 1831 ATATAGTACCAATAGATA-----TACCAGCTATAGTTGTAGTATGTAACACCTGAG 1884
    |||||||
Db 6118 ATATAGTACCAATAGATA-----TACCAGCTATAGTTGTAGTATGTAACACCTGAG 6177

OY 1885 TCATTTACAGAGCCGTGTCCAAAGGTATCCTTTGAGCCAAATCCCATACATTATTGTGCC 1944
    |||||||
Db 6178 TCATTTACAGAGCCGTGTCCAAAGGTATCCTTTGAGCCAAATCCCATACATTATTGTGCC 6237

OY 1945 CGCGTGGTTCCTGATTTTAAATGTAAATTAAGACGTTCAATGGAACAGACCATGTA 2004
    |||||||
Db 6238 CGCGTGGTTCCTGATTTTAAATGTAAATTAAGACGTTCAATGGAACAGACCATGTA 6297

OY 2005 CAAATGTGACAGCATCAATGTACATGGAATCAGGACAGTATGATCAACTCACTGC 2064
    |||||||
Db 6298 CAAATGTGACAGCATCAATGTACATGGAATGGAATTTAGCCAGTATATCAACTCACTGC 6357

OY 2065 TGTTAATGAGCAGTCTAGCAGAGAAGATGTAGTAATTAATGATCTGCCAATTTACAGACA 2124
    |||||||
Db 6358 TGTTAATGAGCAGTCTAGCAGAGAAGATGTAGTAATTAATGATCTGCCAATTTACAGACA 6417

OY 2125 ATGCTAAACCATATATATAGCTGAACACATCTGTGAAATTAATTTGTAAGACCCA 2184
    |||||||
Db 6418 ATGCTAAACCATATATATAGCTGAACCAATCTGTGAAATTAATTTGTAAGACCCA 6477

OY 2185 ACAACAATCAAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGGAGACATTTGTTACAA 2244
    |||||||
Db 6478 ACAACAATCAAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGGAGACATTTGTTACAA 6537

OY 2245 TAGGAAAAATAGGAATATGAGACAAGCACAATTTGTAACATTAGAGCAAAATGGAATG 2304
    |||||||
```

```
Db 6538 TAGGAAAAATAGGAATATGAGACAAGCACAATTTGTAACATTAGTAGACAAAATGGAATA 6597
OY 2305 CCACCTTAAACAGATTAAGCAATTAAGAGAAACATTTTGGAAATTAATAAACAAATAA 2364
    |||||||
Db 6598 ACACCTTAAACAGATTAAGCAATTAAGAGAAACATTTTGGAAATTAATAAACAAATAA 6657

OY 2365 TCTTTAACCAATCCCTCAGAGGGGACCCAGAAAATGTAAACGACAGTTTAAATTTGAGAG 2424
    |||||||
Db 6658 TCTTTAACCAATCCCTCAGAGAGGGACCCAGAAAATGTAAACGACAGTTTAAATTTGAGAG 6717

OY 2425 GGAATTTTCTACTGTATTAATCAACCAACTGTTTAAATAGTACTGTTTAAATAGTACTT 2484
    |||||||
Db 6718 GGAATTTTCTACTGTATTAATCAACCAACTGTTTAAATAGTACTGTTTAAATAGTACTT 6777

OY 2485 GGAGTACTGAAGGGTCAAAATTAACACTGAAGAGTACACAAATCACCCTCAGTACAGAA 2544
    |||||||
Db 6778 GGAGTACTGAAGGGTCAAAATTAACACTGAAGAGTACACAAATCACCCTCAGTACAGAA 6837

OY 2545 TAAACCAATTTATTAACATGTGGCGAGAAAGTAAAGAAAAGCAATGTATGCCCTCCATCA 2604
    |||||||
Db 6838 TAAACCAATTTATTAACATGTGGCGAGAAAGTAAAGAAAAGCAATGTATGCCCTCCATCA 6897

OY 2605 GTGACAATTAATAGATGTTATCAATAATTAATTAAGTGGCTCTATTAAACAAGATGGTGTA 2664
    |||||||
Db 6898 GTGACAATTAATAGATGTTATCAATAATTAATTAAGTGGCTCTATTAAACAAGATGGTGTA 6957

OY 2665 ATTAACAATAATGGGTCCGAGATCTTCAACACTGGAGAGGCGCATATGAGGCAATTTGGA 2724
    |||||||
Db 6958 ATTAACAATAATGGGTCCGAGATCTTCAACACTGGAGAGGCGCATATGAGGCAATTTGGA 7017

OY 2725 GAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2784
    |||||||
Db 7018 GAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7077

OY 2785 AGGCAAGAAGAGAGATGTTGACAGAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2844
    |||||||
Db 7078 AGGCAAGAAGAGAGATGTTGACAGAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7137

OY 2845 TTGGTCTTGTGGAGACAGAGAACACTATATGGCTGTGACCTCAATGACGCTGAGAGGTAC 2904
    |||||||
Db 7138 TTGGTCTTGTGGAGACAGAGAACACTATATGGCTGTGACCTCAATGACGCTGAGAGGTAC 7197

OY 2905 AGGCGAGACAATTAATGCTGATATATGATGAGCAGCAGACAACAATTTGCTGAGAGGCTATTG 2964
    |||||||
Db 7198 AGGCGAGACAATTAATGCTGATATATGATGAGCAGCAGACAACAATTTGCTGAGAGGCTATTG 7257

OY 2965 AGGCGCAACAGACATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGGCTCCAGGCAAGAA 3024
    |||||||
Db 7258 AGGCGCAACAGACATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGGCTCCAGGCAAGAA 7317

OY 3025 TCTGTGGCTGTGGAAGATTAAGTAAAGATCAACAGCTCTCTGGGATTTGGGGTGTCTTG 3084
    |||||||
Db 7318 TCTGTGGCTGTGGAAGATTAAGTAAAGATCAACAGCTCTCTGGGATTTGGGGTGTCTTG 7377

OY 3085 GAAAACATATTGACACACTGCTGTGCTTGGAAATGCTAGTTGAGTAATTAATGCTGCG 3144
    |||||||
Db 7378 GAAAACATATTGACACACTGCTGTGCTTGGAAATGCTAGTTGAGTAATTAATGCTGCG 7437

OY 3145 AACAGATTTGGAATTAACATGACCTGAGTAGAGAGTGGGACAGAGAAATTAACAAATTAACACAA 3204
    |||||||
Db 7438 AACAGATTTGGAATTAACATGACCTGAGTAGAGAGTGGGACAGAGAAATTAACAAATTAACACAA 7497

OY 3205 GCTTAATATCACTCCTTAATTAAGAAATGCAAAACACAGAAAAGAAATGAACAGAAAT 3264
    |||||||
Db 7498 GCTTAATATCACTCCTTAATTAAGAAATGCAAAACACAGAAAAGAAATGAACAGAAAT 7557

OY 3265 TATTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3324
    |||||||
Db 7558 TATTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7617

OY 3325 GGTATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3384
    |||||||
Db 7618 GGTATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7677
```

QY 3385 CTGTACTTCTATAGTAATAGAGTAGGCGAGATATTCACATTATCGTTTCAGACC 3444  
|||||  
Db 7678 CTGTACTTCTGTAGTAGTAATAGAGTAGGCGAGATATTCACATTATCGTTTCAGACC 7737  
QY 3445 ACCTCCCAATCCGAGGGGACCCGACAGGCCGAGAGATAGAGAAAGAGTGGAGAGA 3504  
|||||  
Db 7738 ACCTCCCAATCCGAGGGGACCCGACAGGCCGAGAGATAGAGAAAGAGTGGAGAGA 7797  
QY 3505 GAGACAGAGACAGATCCATTGATTAAGAACGATCCCTTAGCACTTATCTGGAGCATC 3564  
|||||  
Db 7798 GAGACAGAGACAGATCCATTGATTAAGAACGATCCCTTAGCACTTATCTGGAGCATC 7857  
QY 3565 TCGGAGACCTGTGCTCTTACGTACACCGCTTAGAGACTTACTCTTGATTGTAAAGA 3624  
|||||  
Db 7858 TCGGAGACCTGTGCTCTTACGTACACCGCTTAGAGACTTACTCTTGATTGTAAAGA 7917  
QY 3625 GGATTGTGAACTTCTGGAGCGCAGGGGGTGGAGACCCCTCAATATTGGTGAATCTCC 3684  
|||||  
Db 7918 GGATTGTGAACTTCTGGAGCGCAGGGGGTGGAGACCCCTCAATATTGGTGAATCTCC 7977  
QY 3685 TACGTAFTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTGCTCAATGCCACAGCA 3744  
|||||  
Db 7978 TACGTAFTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTGCTCAATGCCACAGCA 8037  
QY 3745 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTACAGACAGCTTATAGACTA 3804  
|||||  
Db 8038 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTACAGAGACCTTATAGACTA 8097  
QY 3805 TTGCGCACATACCTAGAGAAATAAGACAGGGCTTGAAAGGATTTGCTATATAGATGGT 3864  
|||||  
Db 8098 TTGCGCACATACCTAGAGAAATAAGACAGGGCTTGAAAGGATTTGCTATATAGATGGT 8157  
QY 3865 GGCAGAGTGTCAAAAGTATAGTGTATGGATGGCTGTAAAGGAAAGATGACAGA 3924  
|||||  
Db 8158 GGCAGAGTGTCAAAAGTATAGTGTATGGATGGCTGTAAAGGAAAGATGACAGA 8217  
QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGATCTAGA 3971  
|||||  
Db 8218 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCAGTATCTCGAGACCTAGA 8264

Search completed: September 17, 2003, 14:10:44  
Job time : 21922 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 06:21:59 : Search time 1469 Seconds  
(without alignments)  
11446.426 Million cell updates/sec

Title: US-09-913-159A-10

Perfect score: 6229

Sequence: 1 ctgcagcgcctctgacgagc.....attccccgaaagtgcac 6229

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_19Jun03:\*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	6229	100.0	6229	AAA97934	HIV-1 env DNA frag
2	2676.2	43.0	9709	AAO96140	HIV-1 NT4-3 genome
3	2676.2	43.0	9709	AAV81871	Nucleic acid seque
4	2676.2	43.0	9709	AAA97926	HIV-1 env DNA. Hu
5	2676.2	43.0	9709	AAA40298	HIV-1 viral protei
6	2676.2	43.0	9709	AAZ88127	HIV-1 NV5/IIIB rec
7	2676.2	43.0	9709	ABO76897	HIV-1 NT4-3 gp120
8	2676.2	43.0	9709	ABSS7823	HIV-1 genome clone

9	2669.8	42.9	9709	13	AAO22488
10	2658.6	42.7	9965	24	ABO81389
11	2636.6	42.3	15581	18	AAV14355
12	2607.4	41.9	9213	20	AAV74271
13	2599.4	41.7	9213	20	AAV81866
14	2598.2	41.7	3156	7	AAV60128
15	2597.8	41.7	8932	22	AAH76385
16	2597.8	41.7	8933	22	AAF60542
17	2597.8	41.7	8933	22	AAF60543
18	2597.8	41.7	8933	22	AAF60548
19	2597.8	41.7	9425	14	AAO45919
20	2597.8	41.7	9425	14	AAO45922
21	2597.8	41.7	9749	14	AAO45922
22	2596.2	41.7	9213	7	AAV60288
23	2593	41.6	9181	22	AAF24377
24	2593	41.6	9181	25	AAO49652
25	2593	41.6	9181	25	AAO45920
26	2591.4	41.6	9427	14	AAO45920
27	2591.4	41.6	9745	7	AAV60240
28	2590.8	41.6	2945	16	AAV05127
29	2572.2	41.3	9609	24	AAO25515
30	2569.6	41.3	9193	24	AAV49920
31	2561.6	41.1	9193	7	AAV60365
32	2560.4	41.1	2562	22	AAH20891
33	2558.4	41.1	12494	20	AAV83206
34	2558.4	41.1	12494	22	AAV85623
35	2557	41.0	2565	22	AAV56366
36	2543.4	40.8	9088	7	AAV60476
37	2476	39.7	10006	24	AAO25514
38	2454	39.4	2607	7	AAV60077
39	2444.4	39.2	4020	8	AAV71016
40	2333.2	37.5	2641	13	AAO30970
41	2310.4	37.1	2553	13	AAV80949
42	2234.4	35.9	8298	19	AAV18096
43	2206.6	35.4	3806	20	AAO4767
44	2205.6	35.4	3807	18	AAV58551
45	2197.4	35.3	8560	21	AAV3981

#### ALIGNMENTS

RESULT 1  
AAA97934 standard; DNA: 6229 BP.

AAA97934:

26-JAN-2001 (first entry)

HIV-1 env DNA fragment #2 in plasmid pBSCenvATC.

Viral protein; vaccine; anti-viral; anti-HIV; therapy; infection;

env protein; envelope protein; ds.

OS Human immunodeficiency virus type 1.

XX WO200047223-A2.

PN 17-AUG-2000.

PD 03-DEC-1999; 99WO-EP09759.

PE 12-FEB-1999; 99DE-1007485.

PR (STR-A) STRATHMANN & CO AG.

PA Schreiber M;

PI WPI; 2000-549084/50.

XX Viral vaccine comprises a mixture of protein sequence variants of a

DR single viral protein, which is useful for prevention and therapy of

PT

PT viral infections, especially HIV, in humans  
XX  
PS Claim 28; Page 76-78; 79pp; German.  
XX  
CC This invention describes a novel protein vaccine which comprises a  
CC mixture of viral proteins, characterized in that the molecules are  
CC sequence variants of a single viral (partial) protein. The products of  
CC the invention have anti-viral and anti-HIV activity. Mixtures of  
CC structurally different viral proteins, that are sequence variants of a  
CC single protein are useful for production of vaccines for the prevention  
CC and/or therapy of viral infections in humans. The vaccines are especially  
CC useful for prevention and/or therapy of human immunodeficiency virus  
CC (HIV) infection in humans. The vectors and host cells are useful for the  
CC expression of the protein/DNA mixtures, which are also useful for the  
CC preventing and/or therapy of viral infection. This sequence encodes  
CC a fragment of the human immunodeficiency virus (HIV-1) envelope (env)  
CC protein which is used to illustrate the method of the invention.  
XX  
S0 Sequence 6229 BP; 1795 A; 1332 C; 1518 G; 1584 T; 0 other;  
  
Query Match 100.0%; Score 6229; DB 21; Length 6229;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CTGACGGCCCTGTAGCGGCGCATTAAGCGGCGGCGGTGTGTTACGGCAGCGTGA 60  
DB 1 CTGACGGCCCTGTAGCGGCGCATTAAGCGGCGGCGGTGTGTTACGGCAGCGTGA 60  
QY 61 CCGGTACACTTGGCCAGGCGCCCTACGCGCCGCTCTTCGCTTCTCCCTTCCTTTCG 120  
DB 61 CCGGTACACTTGGCCAGGCGCCCTACGCGCCGCTCTTCGCTTCTCCCTTCCTTTCG 120  
QY 121 CCAGCTTCCGCGGCTTCCCGCTCAAGCTTAATCGGCGGCGTCCCTTAAAGTTCGAT 180  
DB 121 CCAGCTTCCGCGGCTTCCCGCTCAAGCTTAATCGGCGGCGTCCCTTAAAGTTCGAT 180  
QY 121 CCAGCTTCCGCGGCTTCCCGCTCAAGCTTAATCGGCGGCGTCCCTTAAAGTTCGAT 180  
DB 121 CCAGCTTCCGCGGCTTCCCGCTCAAGCTTAATCGGCGGCGTCCCTTAAAGTTCGAT 180  
QY 181 TTATGCTTTACGGCACCCTGACCCCAAAAATTGATAGGATGATGTTACAGTATG 240  
DB 181 TTATGCTTTACGGCACCCTGACCCCAAAAATTGATAGGATGATGTTACAGTATG 240  
QY 241 GGCCATGGCCCTGATAGAGGTTTGGCCCTTGGAGTGGAGTCCAGCTTCTTATA 300  
DB 241 GGCCATGGCCCTGATAGAGGTTTGGCCCTTGGAGTGGAGTCCAGCTTCTTATA 300  
QY 301 GTGAGCTCTTGTTCACAACTGGAACAACACTCACTATCTCGTATCTTTGAT 360  
DB 301 GTGAGCTCTTGTTCACAACTGGAACAACACTCACTATCTCGTATCTTTGAT 360  
QY 361 TATAGGAGATTTTCCGATTTGGCTTATGTTAAAAATGAGCTGATTTAACAAAAT 420  
DB 361 TATAGGAGATTTTCCGATTTGGCTTATGTTAAAAATGAGCTGATTTAACAAAAT 420  
QY 421 TTACGGGAATTTTAAACAAATATTAAGCTTACAAATTCATTCGAGCTCAGGCTCG 480  
DB 421 TTACGGGAATTTTAAACAAATATTAAGCTTACAAATTCATTCGAGCTCAGGCTCG 480  
QY 481 CAACCTTTGGGAAGGCGATCGTGGCGCTTCCCTATTTACGAGCTGGGAAG 540  
DB 481 CAACCTTTGGGAAGGCGATCGTGGCGCTTCCCTATTTACGAGCTGGGAAG 540  
QY 541 GGGATGTGCTGCAAGGCGATTTAGTTGGTAACCCAGGCTTTCCAGTACAGAGCTTG 600  
DB 541 GGGATGTGCTGCAAGGCGATTTAGTTGGTAACCCAGGCTTTCCAGTACAGAGCTTG 600  
QY 601 TAAAGAGAGGCGCGTAGGCTAGTATTAATAGTAATCAATTAGGGGTCAATTAGT 660  
DB 601 TAAAGAGAGGCGCGTAGGCTAGTATTAATAGTAATCAATTAGGGGTCAATTAGT 660  
QY 661 CATAGCCCATATATGAGATTCCGGTGTACATTAACGTAATGAGTAAAGGCTG 720  
DB 661 CATAGCCCATATATGAGATTCCGGTGTACATTAACGTAATGAGTAAAGGCTG 720  
QY 721 CCGCCCAAGAGACCCCGCCCATTTGAGCTCAATTAATGAGATGTTCCCATAGTAACGCCA 780

DB 721 CCGCCCAAGAGACCCCGCCCATTTGAGCTCAATTAATGAGATGTTCCCATAGTAACGCCA 780  
QY 781 ATAGGAGCTTCCATTGACGTCATATGGGTGAGATTTAGGTAATCTGCCACTTGGCA 840  
DB 781 ATAGGAGCTTCCATTGACGTCATATGGGTGAGATTTAGGTAATCTGCCACTTGGCA 840  
QY 841 GTACATCAAGGTATTCATATGCAAGTAGCGCCCTTATGACGTCATATGACGTAATG 900  
DB 841 GTACATCAAGGTATTCATATGCAAGTAGCGCCCTTATGACGTCATATGACGTAATG 900  
QY 901 CCCGCTGCGATTAATGCCCATGACATGACCTTATGAGGACTTCTTACCTTGGCAGTATC 960  
DB 901 CCCGCTGCGATTAATGCCCATGACATGACCTTATGAGGACTTCTTACCTTGGCAGTATC 960  
QY 961 TACGTATTAATGATCATGCTTATTTACATGCTGATGCTTTGGCAGTACATCAATGGCGCT 1020  
DB 961 TACGTATTAATGATCATGCTTATTTACATGCTGATGCTTTGGCAGTACATCAATGGCGCT 1020  
QY 1021 GGATAGGCGTTTGTACAGGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCATGAGGAGT 1080  
DB 1021 GGATAGGCGTTTGTACAGGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCATGAGGAGT 1080  
QY 1081 TTGTTTGGCACCAAAATCAACGGGACTTTCACAAATGTGCTAACCACTCCGCCCATTTG 1140  
DB 1081 TTGTTTGGCACCAAAATCAACGGGACTTTCACAAATGTGCTAACCACTCCGCCCATTTG 1140  
QY 1141 ACGCAATGCGGCTAGGCGGTGATGAGGAGTCTATATTAACAGACGCTGTTTATG 1200  
DB 1141 ACGCAATGCGGCTAGGCGGTGATGAGGAGTCTATATTAACAGACGCTGTTTATG 1200  
QY 1201 AACCGTACATCGCCTGAGACGCGCATCAAGCTGTTTACCTCCATAGAGACACCGG 1260  
DB 1201 AACCGTACATCGCCTGAGACGCGCATCAAGCTGTTTACCTCCATAGAGACACCGG 1260  
QY 1261 GACAAATTCGAGCTCGTACCGTCCGACGACCAATGAGAGAGAAATGATACGACT 1320  
DB 1261 GACAAATTCGAGCTCGTACCGTCCGACGACCAATGAGAGAGAAATGATACGACT 1320  
QY 1321 TTGTGAGATGGGGGTGGAATGGGGGACCAATGCTCTTGGGATTTATGATGCTGTG 1380  
DB 1321 TTGTGAGATGGGGGTGGAATGGGGGACCAATGCTCTTGGGATTTATGATGCTGTG 1380  
QY 1381 CTACAGAAAATTTGTGGTCAACGCTCATTTATGGGTAACCTGTTGGGAAGAACCA 1440  
DB 1381 CTACAGAAAATTTGTGGTCAACGCTCATTTATGGGTAACCTGTTGGGAAGAACCA 1440  
QY 1441 CCACCTTATTTTGTGCAATGATGCTTAAGCATATGATGATGAGAGTACATATGTTGGG 1500  
DB 1441 CCACCTTATTTTGTGCAATGATGCTTAAGCATATGATGATGAGAGTACATATGTTGGG 1500  
QY 1501 CCACATGTCCTGTGTCCACAGACGCCCAACCAAGAGAGATGTTGGTAAATGTGA 1560  
DB 1501 CCACATGTCCTGTGTGTCCACAGACGCCCAACCAAGAGAGATGTTGGTAAATGTGA 1560  
QY 1561 CAGAAAATTTTAAACATGTTGGAAGAAATGACATGTTAGAGACGATGATGATATATCA 1620  
DB 1561 CAGAAAATTTTAAACATGTTGGAAGAAATGACATGTTAGAGACGATGATGATATATCA 1620  
QY 1621 GTTTATGGGATCAAGGCTTAAGCATGTGTAATTAATTAACCCACTGTGTTAGTTAA 1680  
DB 1621 GTTTATGGGATCAAGGCTTAAGCATGTGTAATTAATTAACCCACTGTGTTAGTTAA 1680  
QY 1681 AGTGCACTGATTTGAAGAATGATACATTAATACATATGATGAGGAGAAATGATATG 1740  
DB 1681 AGTGCACTGATTTGAAGAATGATACATTAATACATATGATGAGGAGAAATGATATG 1740  
QY 1741 AGAAGAGAGATTAATAAACTGCTTTTCAATATCAGACACAGCATTAAGATTAAGTGC 1800  
DB 1741 AGAAGAGAGATTAATAAACTGCTTTTCAATATCAGACACAGCATTAAGATTAAGTGC 1800  
QY 1801 AGAAGATATGCAATTTCTTTTATTAATGATATGATGATGATGATGATGATGATGATG 1860  
DB 1801 AGAAGATATGCAATTTCTTTTATTAATGATATGATGATGATGATGATGATGATGATG 1860

Db 1801 AGAAGAATATGATCTCTTTTATATAAAGCTGATAGTACCAATAGATATATACAGCTATA 1860  
Oy 1861 GGTTGATAGATTGTAACACCTCATCATTTACACAGGCGCTGCCAAAGGTATATCTTGAGC 1920  
Db 1861 GGTGGAATGATGTAACACCTCATCATTTACACAGGCGCTGCCAAAGGTATATCTTGAGC 1920  
Oy 1921 CAATTCCTCATATTTATGTCGCCGCGCTGTTTGGCATTTCTAAATGATATATATA 1980  
Db 1921 CAATTCCTCATATTTATGTCGCCGCGCTGTTTGGCATTTCTAAATGATATATATA 1980  
Oy 1981 CTTTCATGAGACAGAGCCATGTACAAATGTACACAGTACATATACATGGAATCA 2040  
Db 1981 CGTTCAATGGAACAGAGCCATGTACAAATGTACACAGTACATATGTCACATGGAATCA 2040  
Oy 2041 GGCAGATGATCAACATCAATGCTGTTAAATGGCAGCTAGCAGAGAAAGATGTAGTA 2100  
Db 2041 GGCAGATGATCAACATCAATGCTGTTAAATGGCAGCTAGCAGAGAAAGATGTAGTA 2100  
Oy 2101 TTAGATCTGCCAATTTTCACAGACAAATGCTAAACCATATATAGTACAGCTGAACATCTG 2160  
Db 2101 TTAGATCTGCCAATTTTCACAGACAAATGCTAAACCATATATAGTACAGCTGAACATCTG 2160  
Oy 2161 TAGAAATTAATTTGTACAAAGCCCAACAACAATACAGAAAAAGTATCCGATCCAGAGG 2220  
Db 2161 TAGAAATTAATTTGTACAAAGCCCAACAACAATACAGAAAAAGTATCCGATCCAGAGG 2220  
Oy 2221 GACAGGAGAGCAATTTGTTTCAATAGGAAAAATAGGAAATATAGAACACACATCTGA 2280  
Db 2221 GACAGGAGAGCAATTTGTTTCAATAGGAAAAATAGGAAATATAGAACACACATCTGA 2280  
Oy 2281 ACATTTAGTAGAGCAAAATGGAATGCCACTTTAAAAAGATAGTACGAAATTAAGAGAAC 2340  
Db 2281 ACATTTAGTAGAGCAAAATGGAATGCCACTTTAAAAAGATAGTACGAAATTAAGAGAAC 2340  
Oy 2341 AATTTGGAATTAATTAACAAATTAATCTTTAAAGCAATCTCAGAGGGGAGCCAGAAATG 2400  
Db 2341 AATTTGGAATTAATTAACAAATTAATCTTTAAAGCAATCTCAGAGGGGAGCCAGAAATG 2400  
Oy 2401 TAAAGCAGATTTAATTTGTGAGGGAAATTTTCTACTGTAAATTCACACAACTGTTA 2460  
Db 2401 TAAAGCAGATTTAATTTGTGAGGGAAATTTTCTACTGTAAATTCACACAACTGTTA 2460  
Oy 2461 ATATGACTTGGTTTAATAGTACTGTGAGTACGAAGGTCAAATACACTGAAGAAATG 2520  
Db 2461 ATATGACTTGGTTTAATAGTACTGTGAGTACTGTGAAGGTCAAATACACTGAAGAAATG 2520  
Oy 2521 ACACAAATCACACTCCCATGAGAAATTAACAAATTTATTAACATGTGGCAGGAAGTAGAA 2580  
Db 2521 ACACAAATCACACTCCCATGAGAAATTAACAAATTTATTAACATGTGGCAGGAAGTAGAA 2580  
Oy 2581 AAGCAATGTATGCCCCCTCCCATGAGTACAGTGAACAAATTTAGATGTTCACTGGGC 2640  
Db 2581 AAGCAATGTATGCCCCCTCCCATGAGTACAGTGAACAAATTTAGATGTTCACTGGGC 2640  
Oy 2641 TGTATTAAACAAGAGATGGTGGTAAACAACAAATGGGTCGAGATCTTCAGACTGGAG 2700  
Db 2641 TGTATTAAACAAGAGATGGTGGTAAACAACAAATGGGTCGAGATCTTCAGACTGGAG 2700  
Oy 2701 GAGCGATATGAGGAGCAATTTGAGAGATGATTAATTAATTAATTAAGTAGTAATAATG 2760  
Db 2701 GAGCGATATGAGGAGCAATTTGAGAGATGATTAATTAATTAATTAAGTAGTAATAATG 2760  
Oy 2761 AACCATTTAGAGATGACACCCACGAAGGCAAGAGAGAGTGTGACAGAGAAAAAGAG 2820  
Db 2761 AACCATTTAGAGATGACACCCACGAAGGCAAGAGAGAGTGTGACAGAGAAAAAGAG 2820  
Oy 2821 CAGTGGGAATGAGAGCTTTGTTCTTGGGCTTGGGAGCAGAGAAAGCTATGGGCT 2880  
Db 2821 CAGTGGGAATGAGAGCTTTGTTCTTGGGCTTGGGAGCAGAGAAAGCTATGGGCT 2880  
Oy 2881 GCAGCTCAATGACGCTGACGGTACAGGCGACAGCAATTAATTTGATATATAGTACAGAGC 2940  
Db 2881 GCAGCTCAATGACGCTGACGGTACAGGCGACAGCAATTAATTTGATATATAGTACAGAGC 2940

Oy 2941 AGAACAATTTGCTGAGGGCTATTTAGGGGCAACAGCATCTGTGCAACTCAGAGCTGG 3000  
Db 2941 AGAACAATTTGCTGAGGGCTATTTAGGGGCAACAGCATCTGTGCAACTCAGAGCTGG 3000  
Oy 3001 GCATCAAAACAGCTCCAGCAAGAAATCCGCGCTGTGGAAGAATACCTAAAGATCAACAGC 3060  
Db 3001 GCATCAAAACAGCTCCAGCAAGAAATCCGCGCTGTGGAAGAATACCTAAAGATCAACAGC 3060  
Oy 3061 TCCTGGGGATTTGGGGTTCCTGTGAAAACTATTTGCACACATGCTGTGCTTGGAAATG 3120  
Db 3061 TCCTGGGGATTTGGGGTTCCTGTGAAAACTATTTGCACACATGCTGTGCTTGGAAATG 3120  
Oy 3121 CTAGTTGAGTAATTAATCTGGAACGATTTGGAATGAATGACCTGATGAGAGG 3180  
Db 3121 CTAGTTGAGTAATTAATCTGGAACGATTTGGAATGAATGACCTGATGAGAGG 3180  
Oy 3181 ACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240  
Db 3181 ACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240  
Oy 3241 AGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
Db 3241 AGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
Oy 3301 GGTTAACATTAACAAATTTGGCTGTGTATTAATTAATTAATTAATTAATTAATTAATTA 3360  
Db 3301 GGTTAACATTAACAAATTTGGCTGTGTATTAATTAATTAATTAATTAATTAATTAATTA 3360  
Oy 3361 TGTAGATGTTTAAAGATGATTTTGTGCTGACTTTCATATAGTAATAGATTTAGGAGGAT 3420  
Db 3361 TGTAGATGTTTAAAGATGATTTTGTGCTGACTTTCATATAGTAATAGATTTAGGAGGAT 3420  
Oy 3421 ATTCACATTAATTCGTTTCAGACCCATCCCAATCCGAGGGGAGCCAGAGCCGAG 3480  
Db 3421 ATTCACATTAATTCGTTTCAGACCCATCCCAATCCGAGGGGAGCCAGAGCCGAG 3480  
Oy 3481 GAATGAAGAAGAAGGTGAT 3540  
Db 3481 GAATGAAGAAGAAGGTGAT 3540  
Oy 3541 CCTTAGCATTATCTGGGAGCATCTGCGAGCCCTGTGCTTCAAGTACACACGCTTGA 3600  
Db 3541 CCTTAGCATTATCTGGGAGCATCTGCGAGCCCTGTGCTTCAAGTACACACGCTTGA 3600  
Oy 3601 GAGACTTACTCTTGAATGTAACAGAGATTTGGAACTTCTGGACGACAGGGGTGGAG 3660  
Db 3601 GAGACTTACTCTTGAATGTAACAGAGATTTGGAACTTCTGGACGACAGGGGTGGAG 3660  
Oy 3661 CCCCTCAATATTTGGTGGAAATCTCTACAGTATTTGAGTACAGAACTAAAGAAATAGTCTG 3720  
Db 3661 CCCCTCAATATTTGGTGGAAATCTCTACAGTATTTGAGTACAGAACTAAAGAAATAGTCTG 3720  
Oy 3721 TTAACCTGCTCAATGCCACAGCATATGAGAGAGTGGAGAGAGATGGGTTATAGAAG 3780  
Db 3721 TTAACCTGCTCAATGCCACAGCATATGAGAGAGTGGAGAGAGATGGGTTATAGAAG 3780  
Oy 3781 TATTACAAGAGCTTATAGAGCTATTTGCCACATACCTAGAAGAAATTAAGACAGGCTTGG 3840  
Db 3781 TATTACAAGAGCTTATAGAGCTATTTGCCACATACCTAGAAGAAATTAAGACAGGCTTGG 3840  
Oy 3841 AAAGGATTTTCTATAGATGGTGGCAAGTGTGCAAAAATAGTGTATATGATGGCT 3900  
Db 3841 AAAGGATTTTCTATAGATGGTGGCAAGTGTGCAAAAATAGTGTATATGATGGCT 3900  
Oy 3901 GCTGTAAGGGAAGAATGAGAGAGCTGAGGACAGAGAGATGGGTTGAGAGATCT 3960  
Db 3901 GCTGTAAGGGAAGAATGAGAGAGCTGAGGACAGAGAGATGGGTTGAGAGATCT 3960  
Oy 3961 CGAGATCTAGACTAGAACTAGCTTCGATCCAGACATGATAAGATACATTGATGAGATTGG 4020  
Db 3961 CGAGATCTAGACTAGAACTAGCTTCGATCCAGACATGATAAGATACATTGATGAGATTGG 4020

OY	4021	ACAAACCAACACTAGATGCAGTGAAGAAAAAAATGCTTTATTTGTGTAATTTGTGATGCTAT	4080
Db	4021	ACAAACCAACAACATGAATGCAAGTGAAGAAAAAAATGCTTTATTTGTGTAATTTGTGATGCTAT	4080
OY	4081	TGCTTTATTTGTAAACCTATTAAGCTGCATTAACAAAGTATTAACAACAACTATGCATTTCA	4140
Db	4081	TGCTTTATTTGTAAACCTATTAAGCTGCATTAACAAAGTATTAACAACAACTATGCATTTCA	4140
OY	4141	TTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTTAAAGCAATTAACCTCTTA	4200
Db	4141	TTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTTAAAGCAATTAACCTCTTA	4200
OY	4201	CAATGTGATGATGGCTGATTAATGATTCCTGCCTGCAGGCTTTCGGTGTATGACGGGGAAC	4260
Db	4201	CAATGTGATGATGGCTGATTAATGATTCCTGCCTGCAGGCTTTCGGTGTATGACGGGGAAC	4260
OY	4261	CTGTGACACATGACAGCTCCCGAGACGGTACACAGCTTGTCTGTAAGCGGATGCCGGAGC	4320
Db	4261	CTGTGACACATGACAGCTCCCGAGACGGTACACAGCTTGTCTGTAAGCGGATGCCGGAGC	4320
OY	4321	AGACAACCCGTCAGGGGCGGTGCAAGGGGGTGTGGCGGGTGTCCGGGCGGACGATACAC	4380
Db	4321	AGACAACCCGTCAGGGGCGGTGCAAGGGGGTGTGGCGGGTGTCCGGGCGGACGATACAC	4380
OY	4381	CAGTCACGTAGCGATAGCGAGTGTATCTGGCTTAACCTATGCGGCATCAGAGCAGATTG	4440
Db	4381	CAGTCACGTAGCGATAGCGAGTGTATCTGGCTTAACCTATGCGGCATCAGAGCAGATTG	4440
OY	4441	TACTGAGAGTGCACATATGTCCGGGCGCGGCTGTGCGCTTTTTCATATAGGCTCCGCGCC	4500
Db	4441	TACTGAGAGTGCACATATGTCCGGGCGCGGCTGTGCGCTTTTTCATATAGGCTCCGCGCC	4500
OY	4501	CCTGAGCAGCATCAACAAAAATCGACGGCTCAAGTTCAGAGTGTGGCGAAACCCGACAGACTA	4560
Db	4501	CCTGAGCAGCATCAACAAAAATCGACGGCTCAAGTTCAGAGTGTGGCGAAACCCGACAGACTA	4560
OY	4561	TAAAGATACCAAGCGGTTTCCCGCTGGAAAGCTCCCTCGTGGCTTCCTGTTCGACCCCTG	4620
Db	4561	TAAAGATACCAAGCGGTTTCCCGCTGGAAAGCTCCCTCGTGGCTTCCTGTTCGACCCCTG	4620
OY	4621	CCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGCGGAAGCGTGGCGCTTTCATATAGC	4680
Db	4621	CCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGCGGAAGCGTGGCGCTTTCATATAGC	4680
OY	4681	TCACGCGTATGATATCTCAAGTTCCGGGTAGTGTGCTGCTCCAAAGTCGGGCTGTGTCCAC	4740
Db	4681	TCACGCGTATGATATCTCAAGTTCCGGGTAGTGTGCTGCTCCAAAGTCGGGCTGTGTCCAC	4740
OY	4741	GAACCCCCCGTTCAGCCCCGACCGCTGTGGCCCTTATCCGGTATACATATGCTTTGATGCTCAAC	4800
Db	4741	GAACCCCCCGTTCAGCCCCGACCGCTGTGGCCCTTATCCGGTATACATATGCTTTGATGCTCAAC	4800
OY	4801	CCGGTAAAGACACGACTTATCGCCACTGTGGACAGACCCACTGGTAAACAGATTAGACGAGCG	4860
Db	4801	CCGGTAAAGACACGACTTATCGCCACTGTGGACAGACCCACTGGTAAACAGATTAGACGAGCG	4860
OY	4861	AGGATATGAGGCGGTGTCTACAGAGTTTCTTAAGTGGTGGCCTAACTACCGGCTACACTAGA	4920
Db	4861	AGGATATGAGGCGGTGTCTACAGAGTTTCTTAAGTGGTGGCCTAACTACCGGCTACACTAGA	4920
OY	4921	AGGACAGTATTTGGTATCTCGGCTCTGTGAAGCCATTACCTTCGGAAAAAGAGTTGGT	4980
Db	4921	AGGACAGTATTTGGTATCTCGGCTCTGTGAAGCCATTACCTTCGGAAAAAGAGTTGGT	4980
OY	4981	AGCTCTTGATTCGGCAACAACACCGCTGTGTAGCGGTGTGTTTTTTGTTCGACAGAG	5040
Db	4981	AGCTCTTGATTCGGCAACAACACCGCTGTGTAGCGGTGTGTTTTTTGTTCGACAGAG	5040
OY	5041	CAGATTCAGCGCAAAAAAAGATCTCAAGAGATCTTTGATCTTTTCACGGGGTCT	5100
Db	5041	CAGATTCAGCGCAAAAAAAGATCTCAAGAGATCTTTGATCTTTTCACGGGGTCT	5100
OY	5101	GACGCTCAGTGGAACGAAAACTACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGG	5160

Db	5101	GAGCGCTAGGGAACGAAAACCTCAGCTTAAAGGATTTGGTCATAGATATTCAAAAAGG	5160
OY	5161	ATCTCCACNAGATCCCTTTTAAATTTAAAAATAGATTTTAAATCAATCTAAAGATATAT	5220
Db	5161	ATCTCCACNAGATCCCTTTTAAATTTAAAAATGAAAGTTTAAATCAATCTAAAGATATAT	5220
OY	5221	GAGTAAACTGGTGTGCACAGTTTACCAATGCTTAAATCAGTGAAGCACCTATCTCAGCATC	5280
Db	5221	GAGTAAACCTGGTGTGCAGATTACCAATGCTTAAATCAGTGAAGCACCTATCTCAGCATC	5280
OY	5281	TGTCATTTTGGTTCATCCCATAGTGGCTGACTCCCGTCGCTAGATTAATACATACAGG	5340
Db	5281	TGTCATTTTGGTTCATCCCATAGTGGCTGACTCCCGTCGCTAGATTAATACATACAGG	5340
OY	5341	GAGGCGTTTACCATCTGCGCCCGAGTGTGCTGCAATGATACCGCGAGACCCAGCTCACGGCT	5400
Db	5341	GAGGCGTTTACCATCTGCGCCCGAGTGTGCTGCAATGATACCGCGAGACCCAGCTCACGGCT	5400
OY	5401	CCAGATTTTATCAGCATTAAACAGCGACGCGGAAGGGGCGAGCGCAGAAAGTGGTCTGCA	5460
Db	5401	CCAGATTTTATCAGCATTAAACAGCGACGCGGAAGGGGCGAGCGCAGAAAGTGGTCTGCA	5460
OY	5461	ACCTTATCCGCTCCATCCAGTCTATTATTTGTCCGGGAAGCTAGAGTAAATGATTGCG	5520
Db	5461	ACCTTATCCGCTCCATCCAGTCTATTATTTGTCCGGGAAGCTAGAGTAAATGATTGCG	5520
OY	5521	CCAGTTAATATGTTTGGGCAACGTTGTGGCATTTGCTACAGGATGTTGTGTCAACGCTCG	5580
Db	5521	CCAGTTAATATGTTTGGGCAACGTTGTGGCATTTGCTACAGGATGTTGTGTCAACGCTCG	5580
OY	5581	TGCTTTGGTATVGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGATTACATGATCC	5640
Db	5581	TGCTTTGGTATVGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGATTACATGATCC	5640
OY	5641	CCCATGTTGGCAAAAAAGGGGTAGCTCCCTCGATCCCGATCGTTGTACAGAAATGAG	5700
Db	5641	CCCATGTTGGCAAAAAAGGGGTAGCTCCCTCGATCCCGATCGTTGTGTAGAAATGAG	5700
OY	5701	TTGGCCGCACTGTTATCAGTCATGTTATGAGCAGCACTGCATTAATCTCTACTGTCATG	5760
Db	5701	TTGGCCGCACTGTTATCAGTCATGTTATGAGCAGCACTGCATTAATCTCTACTGTCATG	5760
OY	5761	CCATCCGTAAAGATGCTTTTCTGTGATGCTGTGATGTTACTCAACATCTATTGTGAAATAG	5820
Db	5761	CCATCCGTAAAGATGCTTTTCTGTGATGCTGTGATGTTACTCAACATCTATTGTGAAATAG	5820
OY	5821	TGTTATGGGGGACCGGATGTCCTTCCCGGGGTAAATACGGGATTAATACCGCCACAT	5880
Db	5821	TGTTATGGGGGACCGGATGTCCTTCCCGGGGTCAATACGGGATTAATACCGCCACAT	5880
OY	5881	AGCAGAACTTTAAAGTGCTCATCTATGSAAAAAGCTTCTTGGGGCGAAAACCTCAGAG	5940
Db	5881	AGCAGAACTTTAAAGTGCTCATCTATGSAAAAAGCTTCTTGGGGCGAAAACCTCAGAG	5940
OY	5941	ATCTTACCGGTGTTGAGATCCAGTTGATGTAAACCACCTGTCACCCAACTGATCTTCA	6000
Db	5941	ATCTTACCGGTGTTGAGATCCAGTTGATGTAAACCACCTGTCACCCAACTGATCTTCA	6000
OY	6001	GCAATCTTTTCTTTCACACGAGCTTCTGGGTGAGCAAAAAACGSAAGGCAAAATGCCGA	6060
Db	6001	GCAATCTTTTCTTTCACACGAGCTTCTGGGTGAGCAAAAAACGSAAGGCAAAATGCCGA	6060
OY	6061	AAAAAGGGAATTAAGGGCGACAGCGAAATGTGAATATCTCATCTCTCTTTTCAATAT	6120
Db	6061	AAAAAGGGAATTAAGGGCGACAGCGAAATGTGAATATCTCATCTCTCTTTTCAATAT	6120
OY	6121	TATTAAGCACTTTATCAGGGTTATTTGTCTCATGACGAGATCAATTTTGAATGTATTTAG	6180
Db	6121	TATTAAGCACTTTATCAGGGTTATTTGTCTCATGACGAGATCAATTTTGAATGTATTTAG	6180
OY	6181	AAAAATTAACAAATAGGGGTTCGCGCACATTTCCCGAAAGATGCGAC 6229	

DB	6181	AAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTCCAC	6229
RESULT 2			
AAQ96140			
ID	AAQ96140	standard; DNA: 9709 BP.	
XX			
AC	AAQ96140;		
XX			
DT	24-FEB-1996	(first entry)	
XX			
DE	HIV-1 NL4-3 genomic DNA.		
XX			
KW	HIV-1; AIDS; attenuation; vaccine; ss.		
OS	Human immunodeficiency virus type 1.		
XX			
PN	W09521912-A1.		
XX			
PD	17-AUG-1995.		
XX			
PF	14-FEB-1995;	95WO-AU00063.	
XX			
PR	23-DEC-1994;	94AU-0000284.	
XX			
PR	14-FEB-1994;	94AU-0003864.	
XX			
PR	21-FEB-1994;	94AU-0004002.	
XX			
PA	(AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.		
PA	(MACF-) MACFARLANE BURNET CENT MEDICAL.		
XX			
PI	Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA;		
XX			
DR	WPI; 1995-293115/38.		
XX			
PT	New non-pathogenic HIV-1 strain carrying a deletion in its nef gene		
PT	or LTR region - can be used in a vaccine to inhibit/reduce		
PT	productive infection in an individual by a pathogenic strain		
XX			
PS	Disclosure; Page 8-14; 301pp; English.		
CC	A genomic nucleotide sequence from pathogenic HIV-1 strain NL4-3		
CC	is given in AAQ96140. Attenuated non-pathogenic variants of NL4-3		
CC	are obtd. by deletion of portions of the nef gene (AAQ96141) and/or		
CC	LTR region (see AAQ96406-Q97166).		
XX			
SO	Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other;		
Query Match	43.0%; Score 2676.2; DB 16; Length 9709;		
Best Local Similarity	99.9%; Pred. No. 5.2e-314;		
Matches 2678; Conservative	0; Mismatches 3; Indels 0; Gaps 0;		
QY	1281	CCATGAGAGTAAAGAGAAAGTATCAGCAGCTGTGGAAGTGGGGGTGCAAAATGGGGCACCA	1350
DB	6219	CAATGAGAGTAAAGAGAAAGTATCAGCAGCTGTGGAAGTGGGGGTGCAAAATGGGGCACCA	6278
QY	1351	TGTCCTCTGGGATATTGATGATCTGTAGTGTCTACAGAAAAATTGGGGTCAACGCTTATT	1410
DB	6279	TGTCCTCTGGGATATTGATGATCTGTAGTGTCTACAGAAAAATTGGGGTCAACGCTTATT	6338
QY	1411	ATGGGGTACTGTGTGGAAGGAGCAACACCACTCTATTTTGTGCATCAGATGCTAAG	1470
DB	6339	ATGGGGTACTGTGTGGAAGGAGCAACACCACTCTATTTTGTGCATCAGATGCTAAG	6398
QY	1471	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCA	1530
DB	6399	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCA	6458
QY	1531	ACCCACAGAGTATGTTTGGTAAATGAGACAGAAAAATTTTACATGTTGAAAAATGACA	1590
DB	6459	ACCCACAGAGTATGTTTGGTAAATGAGACAGAAAAATTTTACATGTTGAAAAATGACA	6518
QY	1591	TGTTAGAACAGATGCATGAGATATTAATCACTTTATGGATCAAAAGCCTTAAGCCATGTG	1650

Db	6519	TGTTAGAACAGATGCATGAGCATATTAATCAGTTTATGGATCCAAAGCTTAAGGCATGTG	6578
Qy	1651	TAAATATPACCCCACTGCTGTGTATTAAGTCCACGTATTTGGAATGATPACTATA	1710
Db	6579	TAAATATPACCCCACTGCTGTGTATTAAGTCCACGTATTTGGAATGATPACTATA	6638
Qy	1711	CCAAATAGTAGTAGCGGGGAGAAATGATTAATGAGAAAAGAGAGATATAAAAACCTGCTTTTCA	1770
Db	6639	CCAAATAGTAGTAGCGGGGAGAAATGATTAATGAGAAAAGAGAGATATAAAAACCTGCTTTTCA	6698
Qy	1771	ATATCAGACACAAGCATAGAGATTAAGTGCAGAAAGAAATATGCACTCTTTTATTAACCTTG	1830
Db	6699	ATATCAGACACAAGCATAGAGATTAAGTGCAGAAAGAAATATGCACTCTTTTATTAACCTTG	6758
Qy	1831	ATTATAGTACCAATAGATPATAACCAAGCTATAGTGTATAGTTGTAAACACCTCAGTCATTA	1890
Db	6759	ATTATAGTACCAATAGATPATAACCAAGCTATAGTGTATAGTTGTAAACACCTCAGTCATTA	6818
Qy	1891	CACAGGCTGTCCAAAGGTATCCTTTGAGCCAAATTCCTATACATTAATGTGCCCCGGCTG	1950
Db	6819	CACAGGCTGTCCAAAGGTATCCTTTGAGCCAAATTCCTATACATTAATGTGCCCCGGCTG	6878
Qy	1951	GTTTTGGCATTTCTAAAAATGTAATATAAGACGTTCAATGGAACAGACGACGATGCAATG	2010
Db	6879	GTTTTGGCATTTCTAAAAATGTAATATAAGACGTTCAATGGAACAGACGATGCAATG	6938
Qy	2011	TCAGACAGTACCAATGTACACATGGAATCAGGCCAGTAGATGTAACAATCACTGCTGTAA	2070
Db	6939	TCAGACAGTACCAATGTACACATGGAATCAGGCCAGTAGATGTAACAATCACTGCTGTAA	6998
Qy	2071	ATGCGAGTCTPAGCAGAAAGAGATGTAGTAATTAATGCAATTCGCCAATTTTCCACAGCAATGCTA	2130
Db	6999	ATGCGAGTCTPAGCAGAAAGAGATGTAGTAATTAATGCAATTCGCCAATTTTCCACAGCAATGCTA	7058
Qy	2131	AAACCATATATGTAACAGCTGGAACATCTGTAGAAATTAATTTGTACAGAGCCCAACAACA	2190
Db	7059	AAACCATATATGTAACAGCTGGAACATCTGTAGAAATTAATTTGTACAGAGCCCAACAACA	7118
Qy	2191	ATACAGAGAAAAGTATCCGTATCCAGAGGGGACAGGGAGAGCAATTTGTTTCATPAGGAA	2250
Db	7119	ATACAGAGAAAAGTATCCGTATCCAGAGGGGACAGGGAGAGCAATTTGTTTCATPAGGAA	7178
Qy	2251	AAATAGGAAATATGAGACACACATTTGTAACATTAAGTAGAGCAAAATGGAATGCCACTT	2310
Db	7179	AAATAGGAAATATGAGACACACATTTGTAACATTAAGTAGAGCAAAATGGAATGCCACTT	7238
Qy	2311	TAAACAGATAGCTAGCAAAATTAAGAGACAAATTTGCAATATATAAACAATTAATCTTTA	2370
Db	7239	TAAACAGATAGCTAGCAAAATTAAGAGACAAATTTGCAATATATAAACAATTAATCTTTA	7298
Qy	2371	AGCAATCTCTCAGAGGGGAGCCACAAATTTGAAGGACAGTTTAATTTGAGGAGGGGAT	2430
Db	7299	AGCAATCTCTCAGAGGGGAGCCACAAATTTGAAGGACAGTTTAATTTGAGGAGGGGAT	7358
Qy	2431	TTTTTACTGTAAATTTCAACACACAGCTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTA	2490
Db	7359	TTTTTACTGTAAATTTCAACACACAGCTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTA	7418
Qy	2491	CTGAAGGTCATTAATACACTGAAGAGAGTACACAACTCCATGCAATTAATAAC	2550
Db	7419	CTGAAGGTCATTAATACACTGAAGAGAGTACACAACTCCATGCAATTAATAAC	7478
Qy	2551	AATTTATTAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCCTCCATCAGTGGAC	2610
Db	7479	AATTTATTAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCCTCCATCAGTGGAC	7538
Qy	2611	AAATAGATGTTTCATCAAAATATTAATCTGGGCTGCTATTAAACAAGAGATGGTGGTAATPACA	2670
Db	7539	AAATAGATGTTTCATCAAAATATTAATCTGGGCTGCTATTAAACAAGAGATGGTGGTAATPACA	7598
Qy	2671	ACAATGGTCCGAGATCTTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGGAGAAAGTG	2730
Db	7599	ACAATGGTCCGAGATCTTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGGAGAAAGTG	7658









Db	7179	AAATAGGAAATATATGACACACACACTTGTAACTATTAGACGAAATGGAATGCCACTT	7238
OY	2311	TAAACACGATAGCTAGCAAAATTAAGAGAAACAATTGGAAATATATAAACAATATCTTTA	2370
Db	7239	TAAACAGATATGCTAGCAAAATTAAGAGACAAATTTGGAAATATATAAACAATATCTTTA	7298
OY	2371	AGCAATCCCTCAGAGAGGGGACCCAGAAATTTGTAAACGACAGTTTAAATTTGAGGGGAAAT	2430
Db	7299	AGCAATCCCTCAGAGAGGGGACCCAGAAATTTGAAACGACAGTTTAAATTTGAGGGGAAAT	7358
OY	2431	TTTTCTCTCTGTAATTCAAACACAACCTGTTTAACTACTCTGGTTTAATATAGTACTGGAGTA	2490
Db	7359	TTTTCTCTCTGTAATTCAAACACAACCTGTTTAACTACTCTGGTTTAATATAGTACTGGAGTA	7418
OY	2491	CTGAGAGGCTCAAAATACACTGACAGAGAGTGACACAATCACACTGCCATCGCAATPAAAC	2550
Db	7419	CTGAGAGGCTCAAAATACACTGACAGAGAGTGACACAATCACACTGCCATCGCAATPAAAC	7478
OY	2551	AATTTATPAAACATGTGGCAGAGAGTAGGAAAAGCAATGTATGCCCTCCATCCAGTGGAC	2610
Db	7479	AATTTATPAAACATGTGGCAGAGAGTAGGAAAAGCAATGTATGCCCTCCATCCAGTGGAC	7538
OY	2611	AAATTTAGTGTCTACCAAAATTTTCTCTGGCGCTGATTAACAAGAGATGTGTGTAATACA	2670
Db	7539	AAATTTAGTGTCTACCAAAATTTTCTGGCGCTGATTAACAAGAGATGTGTGTAATACA	7598
OY	2671	ACAATGGGTCCGCAATCTTCAGACCTCGAGAGGCGCATATAGAGGACAAATGGAGAAAGTG	2730
Db	7599	ACAATGGGTCCGCAATCTTCAGACCTCGAGAGGCGCATATAGAGGACAAATGGAGAAAGTG	7658
OY	2731	AATTTATPAAATATPAAAGTAGTTAAAAATTGACCAATTAGAGTAGACCCACCAAGGCAA	2790
Db	7659	AATTTATPAAATATPAAAGTAGTTAAAAATTGACCAATTAGAGTAGACCCACCAAGGCAA	7718
OY	2791	AGACAGAGGTGTGTGTCAGAGAGAAAAAGACAGTGGGAAATGTGAGACTTTGTTCTTGGGT	2850
Db	7719	AGACAGAGGTGTGTGTCAGAGAGAAAAAGACAGTGGGAAATGTGAGACTTTGTTCTTGGGT	7778
OY	2851	TCTTTGGGAGCAGCAGAGAAAGCACTATGTGGCTGTACAGTCAATGACCTGACGGTAGAGGCA	2910
Db	7779	TCTTTGGGAGCAGCAGAGAAAGCACTATGTGGCTGTACAGTCAATGACCTGACGGTAGAGGCA	7838
OY	2911	GACAAATATTTGTCTGATATATAGTGCACGACGAGAAACAATTGCTGAGGGCTATTTGAGCGCG	2970
Db	7839	GACAAATATTTGTCTGATATATAGTGCACGACGAGAAACAATTGCTGAGGGCTATTTGAGCGCG	7898
OY	2971	AACAGCAATCTGTGCACACTGCACACTCTGTGGGCAATCAAAACGCTCAGGCAAGAAATCCTGG	3030
Db	7899	AACAGCAATCTGTGTCCAACTGCACACTCTGTGGGCAATCAAAACGCTCAGGCAAGAAATCCTGG	7958
OY	3031	CTGTGGAAGATATACCTTAAGAGATCAACAGCTCTGTGGGATTTTGGGTTTGCTCTGCAAAAC	3090
Db	7959	CTGTGGAAGATATACCTTAAGAGATCAACAGCTCTGTGGGATTTTGGGTTTGCTCTGCAAAAC	8018
OY	3091	TCATTTGCACCACTGCTGTGTGCTTTGGAAATGCTAATTAATTTCTGTGAAACAGA	3150
Db	8019	TCATTTGCACCACTGCTGTGTGCTTTGGAAATGCTAATTAATTTCTGTGAAACAGA	8078
OY	3151	TTTGGAAATPAAACATATACCTCGATGTGAGATGTGAGATGCTAATTAATTTCTGTGAAACAGA	3210
Db	8079	TTTGGAAATPAAACATATACCTCGATGTGAGATGTGAGATGCTAATTAATTTCTGTGAAACAGA	8138
OY	3211	TACACCTCCCTTAATTGAGAGAAATCGGAAAAACCGACAAGAAAAAATGAACACAATTTATTTGG	3270
Db	8139	TACACCTCCCTTAATTGAGAGAAATCGGAAAAACCGACAAGAAAAAATGAACACAATTTATTTGG	8198
OY	3271	AATTAGATPAAATGGGCAAGTTTGTGGAATTTGTTTAAACATPACAAATTTGGCTGTGATATA	3330
Db	8199	AATTAGATPAAATGGGCAAGTTTGTGGAATTTGTTTAAACATPACAAATTTGGCTGTGATATA	8258
OY	3331	TAAAAATTTATTCATATATGATATAGTAGAGGCTTTGTAGAGTTTAAAGAAATAGTTTTCCTGTAC	3390
Db	8259	TAAAAATTTATTCATATATGATATAGTAGAGGCTTTGTAGAGTTTAAAGAAATAGTTTTCCTGTAC	8318

QY	3331	TTTCTATATGTAATATAGACTTGGCAGGGGATATTACCATTTATCGTTTCAGACCCACTCC	3450
Db	8319	TTTTCTATAGTCAATATAGAGTTTAGCGAGGATATTACCATTTATCGTTTCAGACCCACTCC	8378
QY	3451	CAATCCGAGGGGACCCGACAGGGCCCGAAGGAATAGAAGAAAGAGTGGAGAGAGACACA	3510
Db	8379	CAATCCGAGGGGACCCGACAGGGCCCGAAGGAATAGAAGAAAGAGTGGAGAGAGACACA	8438
QY	3511	GAGACAGATCCATTCGATTAGTGAACGAGATCCTTAGCACTTATCTGGGAGCATCTGGCGA	3570
Db	8439	GAGACAGATCCATTCGATTAGTGAACGAGATCCTTAGCACTTATCTGGGAGCATCTGGCGA	8498
QY	3571	GCCCTGTGCTCTTACGCTACCTACCGCTTGAGAGACTTACTCTGATTTGAACAGAGATTG	3630
Db	8499	GCCCTGTGCTCTTACGCTACCTACCGCTTGAGAGACTTACTCTGATTTGAACAGAGATTG	8558
QY	3631	TGGAACCTCTGGAGCGCAGGGGGGTGGGAAGCCCAATATATTTGGTGGAAATCCTACAGT	3690
Db	8559	TGGAACCTCTGGAGCGCAGGGGGGTGGGAAGCCCAATATATTTGGTGGAAATCCTACAGT	8618
QY	3691	ATTGAGTCTAGGAACCTAAAGAAATAGTGTCTTTAACTTGCTCAATGCCACAGCCATPACG	3750
Db	8619	ATTGAGTCTAGGAACCTAAAGAAATAGTGTCTTTAACTTGCTCAATGCCACAGCCATPACG	8678
QY	3751	TAGCTGAGGGGACAGATATAGGTTTATAGAAATATTACAAGAGCTTATATAGACTATTGGCC	3810
Db	8679	TAGCTGAGGGGACAGATATAGGTTTATAGAAATATTACAAGAGCTTATATAGACTATTGGCC	8738
QY	3811	ACATACCTAGTAAGATATAAGACAGGCTTGGAAAGATTTGCTATAGATGGTGGCGAAG	3870
Db	8739	ACATACCTAGTAAGATATAAGACAGGCTTGGAAAGATTTGCTATAGATGGTGGCGAAG	8798
QY	3871	TGCTCAAAAAGTACTGTGATTTGGATGGCTGCTCTTAAGGGAAGATGAGACGACTGAG	3930
Db	8799	TGCTCAAAAAGTACTGTGATTTGGATGGCTGCTCTTAAGGGAAGATGAGACGACTGAG	8858
QY	3931	CCAGCAGCAGATGGGGTGGCGAGCATTTCTCGAATCTAGA 3971	
Db	8859	CCAGCAGCAGATGGGGTGGCGAGCATTTCTCGAATCTAGA 8899	
RESULT 5			
AAAA0298			
ID	AAAA0298 standard; DNA; 9709 BP.		
AC	AAA40298:		
XX			
DT	02-NOV-2000 (first entry)		
XX			
DE	HIV-1 viral protein operon encoding polynucleotide.		
XX			
KW	Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;		
KW	gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;		
KW	carcinoma; melanoma; env gene; viral protein; ds.		
XX			
OS	Human immunodeficiency virus type 1.		
XX			
PN	EP1006196-A2.		
XX			
PD	07-JUN-2000.		
XX			
PF	25-NOV-1999; 99EP-0250415.		
XX			
PR	26-NOV-1998; 98DE-1056463.		
XX			
PA	(PETT-) PETTE INST HEINRICH.		
XX			
PI	Von Laer MD;		
XX			
WP1	2000-378268/33.		
DR	P-PSDB; AAB10046, AAB10047, AAB10048, AAB10049, AAB10050, AAB10051,		
DR	AAB10052, AAB10053, AAB10054.		



```

Db      8019 TCATTTGACCACTGCTGCTGCTTGAGATGCTAGTGGAGATTAATCTCTGGACACGA 8078
OY      3151 TTTCGAATTAACATGACCTGATGAGTGGGACAGAGAAATTAACAATTACACAACTTAA 3210
Db      8079 TTTCGAATTAACATGACCTGATGAGTGGGACAGAGAAATTAACAATTACACAACTTAA 8138
OY      3211 TACACTGCTTAATTAAGAAATCGCAAAACCAAGAAAGAAATTAACAATTATTTGG 3270
Db      8139 TACACTGCTTAATTAAGAAATCGCAAAACCAAGAAAGAAATTAACAATTATTTGG 8198
OY      3271 AATTAGATAATGGGCAAGTTTGTGAATTTGTTTAACATTAACAATTTGGCTGTATTA 3330
Db      8199 AATTAGATAATGGGCAAGTTTGTGAATTTGTTTAACATTAACAATTTGGCTGTATTA 8258
OY      3331 TAAATTAATTAATTAATGATAGAGAGCTTGTAGTTTAAGAAATTAATTTTGGCTGAC 3390
Db      8259 TAAATTAATTAATTAATGATAGAGAGCTTGTAGTTTAAGAAATTAATTTTGGCTGAC 8318
OY      3391 TTTCATATGTAATTAAGATTAAGGAGGATTAATTCACCTTATTCCTTTGAGACCCCTCC 3450
Db      8319 TTTCATATGTAATTAAGATTAAGGAGGATTAATTCACCTTATTCCTTTGAGACCCCTCC 8378
OY      3451 CAATCCGAGGGGACCCGACAGGCCCGAAGAAATAGAGAAAGGTGGAGAGAGACA 3510
Db      8379 CAATCCGAGGGGACCCGACAGGCCCGAAGAAATAGAGAAAGGTGGAGAGAGACA 8438
OY      3511 GAGACAGATTCATTCGATTAAGTAAAGGATTCCTTAAGCACTTAATCTGAGAGATCTGCGGA 3570
Db      8439 GAGACAGATTCATTCGATTAAGTAAAGGATTCCTTAAGCACTTAATCTGAGAGATCTGCGGA 8498
OY      3571 GCCTGTGCTCTTCAGTACACCGCTTGAGAGATTCCTTGAATTTGAACGAGATTTG 3630
Db      8499 GCCTGTGCTCTTCAGTACACCGCTTGAGAGATTCCTTGAATTTGAACGAGATTTG 8558
OY      3631 TGGAACTTCTGGAGACGAGGGGTGGGAAGCCCTCAAAATTTGTTGGAATCTCTACAGT 3690
Db      8559 TGGAACTTCTGGAGACGAGGGGTGGGAAGCCCTCAAAATTTGTTGGAATCTCTACAGT 8618
OY      3691 ATTGAGAGTCAGAACTAAAGAAATAGTCTGTTAACTTCTCAATGCCACAGCATATGACG 3750
Db      8619 ATTGAGAGTCAGAACTAAAGAAATAGTCTGTTAACTTCTCAATGCCACAGCATATGACG 8678
OY      3751 TAGCTGAGGGGACGATAGGGGTTTAAAGTATTACAAGCAGCTATAGAGCTATTTGCC 3810
Db      8679 TAGCTGAGGGGACGATAGGGGTTTAAAGTATTACAAGCAGCTATAGAGCTATTTGCC 8738
OY      3811 ACATACCTAGAAGATAAGACAGAGGCTTGAAGGATTTTGTATTAAGATGGTGGCAAG 3870
Db      8739 ACATACCTAGAAGATAAGACAGAGGCTTGAAGGATTTTGTATTAAGATGGTGGCAAG 8798
OY      3871 TGGTCAAAAAGTAGTGTATTTGGATGGCTCTGTAAGGAAAGATTAAGATGGTGGCAAG 3930
Db      8799 TGGTCAAAAAGTAGTGTATTTGGATGGCTCTGTAAGGAAAGATTAAGATGGTGGCAAG 8858
OY      3931 CCAGAGCAGATGGGGTGGGAGCAGTATCTGAAGATCTAGA 3971
Db      8859 CCAGAGCAGATGGGGTGGGAGCAGTATCTGAAGCCTAGA 8899

```

RESULT 6  
AAZ88127  
ID AAZ88127 standard; DNA; 9709 BP.

AC AAZ88127;  
XX  
DT 20-APR-2000 (first entry)  
XX  
DE HIV-1 Nf5/IIIB recombinant clone pNL4-3, complete genome.  
XX  
KW LentiViral vector; packaging; gag; pol; gene therapy; infection;  
KW gene expression; PCR primer; ss.  
XX

```

OS      Human immunodeficiency virus type 1.
OS      Synthetic.
PN      WO200000600-A2.
XX      PD 06-JAN-2000.
XX      PF 26-MAY-1999; 99WO-US11516.
XX      PR 26-MAY-1998; 98US-0086635.
XX      PA (CHAN/) CHANG L.
PI      Chang L.
DR      WPI; 2000-137067/12.
XX      PT New packaging vector comprising a nucleotide sequence encoding Gag and
PT      Pol proteins of a reference lentivirus useful for the delivery of
PT      non-lentiviral genes to target cells.
XX      PS Example; Page 258-263; 311pp; English.
XX      CC The present invention describes a packaging vector (PV) comprising a
CC      nucleotide sequence encoding Gag and Pol proteins of a reference
CC      lentivirus that differs from the reference lentivirus at least in that:
CC      (a) its major splice donor site is either deleted or is insufficiently
CC      different from the reference lentivirus so that it is not a potential
CC      site for homologous recombination; and (b) it lacks a functional major
CC      packaging signal so that the introduced vector causes the host cell to
CC      produce packaging vector particles comprising functional Gag and Pol
CC      proteins. The vectors are useful for transforming (eukaryotic) cells to
CC      express specific genes at high levels, e.g. for gene therapy. The
CC      improved vectors are safer, yet permit increased efficiency of packaging
CC      the recombinant viral genome and increased long-term gene expression.
CC      These properties are required for gene therapy as a means of treating
CC      infectious and non-infectious diseases. Unlike other retroviruses, the
CC      lentiviruses are able to infect non-dividing cells. The present sequence
CC      represents the HIV-1 Nf5/IIIB recombinant clone pNL4-3, complete genome
CC      which is given in the exemplification of the present invention.
XX      SX Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other.

```

Query Match 43.0%; Score 2676.2; DB 21; Length 9709;  
Best Local Similarity 99.9%; Pred. No. 5.2e-314;  
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY      1291 CCATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGTGGAAATGGGACCA 1350
Db      6219 CAATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGTGGAAATGGGACCA 6278
OY      1351 TGCCTCTGGGATATTTGATGATCTGTAGTGTACAGAAATTTGTGGGTCAACCGTCTATT 1410
Db      6279 TGCCTCTGGGATATTTGATGATCTGTAGTGTACAGAAATTTGTGGGTCAACCGTCTATT 6338
OY      1411 ATGGGATACCTGTGTGGGAAGAGCAACACACACTCTATTTTGTGATCAGATGCTAAG 1470
Db      6339 ATGGGATACCTGTGTGGGAAGAGCAACACACACTCTATTTTGTGATCAGATGCTAAG 6398
OY      1471 CATATGATACAGAGGTACATATATTTTGGGACACATAGCCGTGTACCCACAGACCCA 1530
Db      6399 CATATGATACAGAGGTACATATATTTTGGGACACATAGCCGTGTACCCACAGACCCA 6458
OY      1531 ACCCACAAGAGTAGTATTTGTAATGTGACAGAAATTTTAACATGTGAAATTAAGCA 1590
Db      6459 ACCCACAAGAGTAGTATTTGTAATGTGACAGAAATTTTAACATGTGAAATTAAGCA 6518
OY      1591 TGGTGAACAGATGATGAGATATTAATCAGTTTATGGGATCAAAAGCCTAAAGCATGTG 1650
Db      6519 TGGTGAACAGATGATGAGATATTAATCAGTTTATGGGATCAAAAGCCTAAAGCATGTG 6578
OY      1651 TAAATTAACCCCACTGCTGTAGTTAAAGTGACATGATTTGAAGATGATGATTAATA 1710

```

Db 6579 TAAATTAACCCACTCTGTGTAGTTTAAAGTCACATGATTGTAAGATGATATA 6638  
Oy 1711 CCAATAGTAGAGGGGAGATGATTAATGCAAAAGGAGATATAAAACAGTCTTTCA 1770  
Db 6639 CCAATAGTAGAGGGGAGATGATTAATGCAAAAGGAGATATAAAACAGTCTTTCA 6658  
Oy 1771 ATATCAGACACAAGCATAGAGATTAAGGTGCAGAAAAGATATGATCTTTTAACTTG 1830  
Db 6659 ATATCAGACACAAGCATAGAGATTAAGGTGCAGAAAAGATATGATCTTTTAACTTG 6758  
Oy 1831 ATATAGTACCAATAGATTAATCCAGCTATAGTTGATTAAGTTGTAACCTCAGCTATTA 1890  
Db 6759 ATATAGTACCAATAGATTAATCCAGCTATAGTTGATTAAGTTGTAACCTCAGCTATTA 6818  
Oy 1891 CACAGGCTGTCCAAAGGTATCTTTGAGCCATTTCCATCATTTATTTGCCCCGGCTG 1950  
Db 6819 CACAGGCTGTCCAAAGGTATCTTTGAGCCATTTCCATCATTTATTTGCCCCGGCTG 6878  
Oy 1951 GTTTTGGGATTCCTAAATGTAAATTAATTAAGACGTTCAATGGAACAGACCATGTACAAATG 2010  
Db 6879 GTTTTGGGATTCCTAAATGTAAATTAATTAAGACGTTCAATGGAACAGACCATGTACAAATG 6938  
Oy 2011 TCACACAGTACATGTACACATGGAATCAGGCCAGTATGATCAACTCAACTGCTGTTAA 2070  
Db 6939 TCACACAGTACATGTACACATGGAATCAGGCCAGTATGATCAACTCAACTGCTGTTAA 6998  
Oy 2071 ATGACAGCTAGCAGAAAGATGTAGTAATTAATGATCTGCCATTTTCACAGCAATGCTTA 2130  
Db 6999 ATGACAGCTAGCAGAAAGATGTAGTAATTAATGATCTGCCATTTTCACAGCAATGCTTA 7058  
Oy 2131 AAACCATTAATGTACAGCTGACACATCTGTGAATTAATTTGTAACAGACCCAAACCA 2190  
Db 7059 AAACCATTAATGTACAGCTGACACATCTGTGAATTAATTTGTAACAGACCCAAACCA 7118  
Oy 2191 ATACAAGAAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCATTTGTTACATATAGAA 2250  
Db 7119 ATACAAGAAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCATTTGTTACATATAGAA 7178  
Oy 2251 AAATAGAGAAATATGAGACAGCACATTTGTAACTTAGTAGAGCAAAATGGAATCCACTT 2310  
Db 7179 AAATAGAGAAATATGAGACAGCACATTTGTAACTTAGTAGAGCAAAATGGAATCCACTT 7238  
Oy 2311 TAAACAGATAGCTAGCAAAATTAAGAACAAATTTGCAAAATTAATTAACATATCTTTA 2370  
Db 7239 TAAACAGATAGCTAGCAAAATTAAGAACAAATTTGCAAAATTAATTAACATATCTTTA 7298  
Oy 2371 AGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAGCAGAGTTTAAATTTGAGGGGAGAT 2430  
Db 7299 AGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAGCAGAGTTTAAATTTGAGGGGAGAT 7358  
Oy 2431 TTTTCTACTGTAAATTCACACACACTGTTTAATAGTACTGTGTTAAATAGTACTTGAGTA 2490  
Db 7359 TTTTCTACTGTAAATTCACACACACTGTTTAATAGTACTGTGTTAAATAGTACTTGAGTA 7418  
Oy 2491 CTGGAAGGTCAAAATTAACACTGAGAGAGTGAACAATTCACACTCCCATGCAATTAAC 2550  
Db 7419 CTGGAAGGTCAAAATTAACACTGAGAGAGTGAACAATTCACACTCCCATGCAATTAAC 7478  
Oy 2551 AATTTAATCAATGTGCGAGAGATAGAGAAAGCAATGTATGCCCTCCCATCTGCTGAGAC 2610  
Db 7479 AATTTAATCAATGTGCGAGAGATAGAGAAAGCAATGTATGCCCTCCCATCTGCTGAGAC 7538  
Oy 2611 AATTTAGATGTTTCATCAAAATTAATTAAGTGGCTGCTATTAAAGAGATGTTGTTAATACA 2670  
Db 7539 AATTTAGATGTTTCATCAAAATTAATTAAGTGGCTGCTATTAAAGAGATGTTGTTAATACA 7598  
Oy 2671 ACAATGGGTCGGAGATCTTACAGACTTGAGAGAGCGATATGAGGACAAATTTGGAAGTG 2730  
Db 7599 ACAATGGGTCGGAGATCTTACAGACTTGAGAGAGCGATATGAGGACAAATTTGGAAGTG 7658  
Oy 2731 AATTTAATTAATTAAGTGAATTAAGTGAACCATTAAGAGTGAAGTGAAGTGAAGTGAAG 2790  
Db 7659 AATTTAATTAATTAAGTGAATTAAGTGAACCATTAAGAGTGAAGTGAAGTGAAGTGAAG 7718

Oy 2791 AGAGAGATGTCGACAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTCTTGGGT 2850  
Db 7719 AGAGAGATGTCGACAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTCTTGGGT 7778  
Oy 2851 TCTTGGGAGCAGAGGAAGCACTATGGGCTGCACGTCAATGACGCTGACGGTACAGGCCA 2910  
Db 7779 TCTTGGGAGCAGAGGAAGCACTATGGGCTGCACGTCAATGACGCTGACGGTACAGGCCA 7838  
Oy 2911 GACAAATTTGTCGATTAATGTCAGAGCAAGCAATTTGCTGAGGGCTATTAGAGCGC 2970  
Db 7839 GACAAATTTGTCGATTAATGTCAGAGCAAGCAATTTGCTGAGGGCTATTAGAGCGC 7898  
Oy 2971 AACAGCATCTCTTTCACATCAGACTGTGGGCAATCAACAGCTCAGAGCAAGATCTCTGG 3030  
Db 7899 AACAGCATCTCTTTCACATCAGACTGTGGGCAATCAACAGCTCAGAGCAAGATCTCTGG 7958  
Oy 3031 CTGTGAAAGATTAACCTAAAGGATCAACAGCTCTGTGGGATTTGGGTTGCTCTGGAAAC 3090  
Db 7959 CTGTGAAAGATTAACCTAAAGGATCAACAGCTCTGTGGGATTTGGGTTGCTCTGGAAAC 8018  
Oy 3091 TCATTTGACACACTGCTGTGCTGGAATGCTAGTTGAGTAATTAATCTCTGGAACAGA 3150  
Db 8019 TCATTTGACACACTGCTGTGCTGGAATGCTAGTTGAGTAATTAATCTCTGGAACAGA 8078  
Oy 3151 TTTTGAATTAACATGACCTGATGAGTGGGAGCAGAGAAATTAACAAATTAACAAGCTTAA 3210  
Db 8079 TTTTGAATTAACATGACCTGATGAGTGGGAGCAGAGAAATTAACAAATTAACAAGCTTAA 8138  
Oy 3211 TACACTCCTTAATTTGAAGATCCGAAACCAGACAGAAAAAGATGAACAAGATTTATGG 3270  
Db 8139 TACACTCCTTAATTTGAAGATCCGAAACCAGACAGAAAAAGATGAACAAGATTTATGG 8198  
Oy 3271 AATTAGAATTAATGGGCAAGTTTGTGAATTTGTTAATCAATTAACAAATTTGGCTGTGATA 3330  
Db 8199 AATTAGAATTAATGGGCAAGTTTGTGAATTTGTTAATCAATTAACAAATTTGGCTGTGATA 8258  
Oy 3331 TAAATTAATTCATTAATGATAGTAGAGAGCTTGTGTAAGATAGTTTGTGCTGTAC 3390  
Db 8259 TAAATTAATTCATTAATGATAGTAGAGAGCTTGTGTAAGATAGTTTGTGCTGTAC 8318  
Oy 3391 TTTTCTAATAGGAATAGAGTATAGCAGAGGATTAATCCATTAATGCTTACACACCACCTCC 3450  
Db 8319 TTTTCTAATAGGAATAGAGTATAGCAGAGGATTAATCCATTAATGCTTACACACCACCTCC 8378  
Oy 3451 CAATCCCGAGGGGAGCCGACAGAGCCCGAAGAAATAGAAAGAAAGGTGAGAGAGAGACA 3510  
Db 8379 CAATCCCGAGGGGAGCCGACAGAGCCCGAAGAAATAGAAAGAAAGGTGAGAGAGAGACA 8438  
Oy 3511 GAGACAGATCCATTTGATAGTGAACGGAATCTTAGCACTTAATCTGGGACGATCTGGGA 3570  
Db 8439 GAGACAGATCCATTTGATAGTGAACGGAATCTTAGCACTTAATCTGGGACGATCTGGGA 8498  
Oy 3571 GCCCTGCTCTTACAGTACACACGCTTGAAGACACTTACCTGATTTGTAACAGAGATTG 3630  
Db 8499 GCCCTGCTCTTACAGTACACACGCTTGAAGACACTTACCTGATTTGTAACAGAGATTG 8558  
Oy 3631 TGGAACTTCTGGGAGCGAGGGGGTGGGAACCCCAATATTTGGTGAATCTCCTACAGT 3690  
Db 8559 TGGAACTTCTGGGAGCGAGGGGGTGGGAACCCCAATATTTGGTGAATCTCCTACAGT 8618  
Oy 3691 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGAATGCCACAGCCATAGCAG 3750  
Db 8619 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGAATGCCACAGCCATAGCAG 8678  
Oy 3751 TAGCTGAGGGGACAGATAGGGTTATAGAAATTAACAGACGCTTATAGAGCTATTGCGC 3810  
Db 8679 TAGCTGAGGGGACAGATAGGGTTATAGAAATTAACAGACGCTTATAGAGCTATTGCGC 8738  
Oy 3811 ACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTGCTATAAGATGGGTGGCAG 3870  
Db 8739 ACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTGCTATAAGATGGGTGGCAG 8798





```
|||||
Db AGCAATCCTCAGAGGGAGCCAGAAATTGTAACGCACAGTCTTAAATTGGAGGGGAAT 7358
QY TTTTCACTGTAAATTCACACACACTGTTTATAGTACTTGGTTAATAGACTTGGGTA 2431
Db TTTTCACTGTAAATTCACACACACTGTTTATAGTACTTGGTTAATAGACTTGGGTA 7359
QY CTGAGGGTCAAAATTAACACTGAAGAGTAGACAAATCACACTCCATGCGAATAAAC 2491
Db CTGAGGGTCAAAATTAACACTGAAGAGTAGACAAATCACACTCCATGCGAATAAAC 7419
QY AATTATTAACATGTGGCAGAGAGTGAAGAAAGCAATGTATGCCCTCCATCAGTGCAC 2551
Db AATTATTAACATGTGGCAGAGAGTGAAGAAAGCAATGTATGCCCTCCATCAGTGCAC 7479
QY AATTAGATGTTCATCAAAATATTACTGGGCTGATTAAACAAGAGATGGTATATACA 2611
Db AATTAGATGTTCATCAAAATATTACTGGGCTGATTAAACAAGAGATGGTATATACA 7539
QY ACAATGGGTCGAGATCTTCAGACCTTGAGAGAGCGATATGAGGGACAATTGGAGAGTG 2671
Db ACAATGGGTCGAGATCTTCAGACCTTGAGAGAGCGATATGAGGGACAATTGGAGAGTG 7599
QY ACATGGGTCGAGATCTTCAGACCTTGAGAGAGCGATATGAGGGACAATTGGAGAGTG 2731
Db AATTATTAATTAATTAAGTAACTTAAATTTGACATTTAGAGATGACACCAAGCA 2791
QY AGAGAGAGTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2791
Db AGAGAGAGTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7719
QY TCTTGGGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2851
Db TCTTGGGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7779
QY GACATATATGTCTGATATATAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2911
Db GACATATATGTCTGATATATAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7839
QY AACGACATCTGTTGCAACTCAGACTCTGGGGCATCAACAGCTCCAGGCAAGATCCTGG 2971
Db AACGACATCTGTTGCAACTCAGACTCTGGGGCATCAACAGCTCCAGGCAAGATCCTGG 7899
QY CTGTGAAGAGATACCTAAAGAGATCAACAGCTCTGGGGATTTGGGGTGTCTGGAAAC 3031
Db CTGTGAAGAGATACCTAAAGAGATCAACAGCTCTGGGGATTTGGGGTGTCTGGAAAC 7959
QY TCAATTTGCACCACTGCTGCTGCTGGAATGCTAGTTGAGATTAATTAATCTCTGGAACGA 8019
Db TCAATTTGCACCACTGCTGCTGCTGGAATGCTAGTTGAGATTAATTAATCTCTGGAACGA 8078
QY TTTGGATTAACATGACCTGGATGGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 3151
Db TTTGGATTAACATGACCTGGATGGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 8079
QY TACACTCTCTTAATTGAAGAGATCGCAAAACAGCAGAGAGAGAGAGAGAGAGATTTGG 3211
Db TACACTCTCTTAATTGAAGAGATCGCAAAACAGCAGAGAGAGAGAGAGAGAGATTTGG 8139
QY TACACTCTCTTAATTGAAGAGATCGCAAAACAGCAGAGAGAGAGAGAGAGAGATTTGG 3271
Db TACACTCTCTTAATTGAAGAGATCGCAAAACAGCAGAGAGAGAGAGAGAGAGATTTGG 3330
QY AATTAGATTAATGGGCAAGTTTGTGAATTGGTTTAAATTAACAAATTTGGCTGTGATA 3371
Db AATTAGATTAATGGGCAAGTTTGTGAATTGGTTTAAATTAACAAATTTGGCTGTGATA 8199
QY TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3331
Db TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8259
QY TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 8319
Db TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 8391
QY TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 8319
Db TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 8391
QY CAATCCGAGGGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3451
Db CAATCCGAGGGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3510
```

```
Db CAATCCGAGGGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8379
QY GAGACAGATCCATTCGATTGATGACGAGATCCTTAGACACTTATCTGGAGATCTGCGG 3511
Db GAGACAGATCCATTCGATTGATGACGAGATCCTTAGACACTTATCTGGAGATCTGCGG 8439
QY GCCTGTGCTCTTCACTACACACCGCTTGAGAGACTTACTCTTGAATTTGAAGAGATTG 3571
Db GCCTGTGCTCTTCACTACACACCGCTTGAGAGACTTACTCTTGAATTTGAAGAGATTG 8499
QY TGGAACTTCTGGAGCGAGGGGGTGGGAACCCCTCAAAATTTGTGGAAATCTCTACAG 3631
Db TGGAACTTCTGGAGCGAGGGGGTGGGAACCCCTCAAAATTTGTGGAAATCTCTACAG 8559
QY ATTGAGATCAGAACTTAAGAAATAGTCTTAACTTGTCTCAATGCGACAGCATAGCAG 3691
Db ATTGAGATCAGAACTTAAGAAATAGTCTTAACTTGTCTCAATGCGACAGCATAGCAG 8619
QY TAGCTGAGGGAGCAGATAGGGGTTATAGAAATTTAGAACAGCCTTATAGAGTATTGCGC 3751
Db TAGCTGAGGGAGCAGATAGGGGTTATAGAAATTTAGAACAGCCTTATAGAGTATTGCGC 8679
QY ACATFACTAAGAAATTAAGACAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3811
Db ACATFACTAAGAAATTAAGACAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 8739
QY TGGTCAAAAAGTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 3871
Db TGGTCAAAAAGTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 8799
QY CCAGCAGCAGATGGGGTGGGAGAGATATCTCGAGATCTAGA 3971
Db CCAGCAGCAGATGGGGTGGGAGAGATATCTCGAGATCTAGA 8859

RESULT 8
ABS57823 standard; DNA: 9709 BP.
ID ABS57823
AC ABS57823;
XX 03-FEB-2003 (first entry)
DT 03-FEB-2003 (first entry)
DE HIV-1 genome clone pNL4-3.
XX
XX HIV-1; ss: genome; retroviral vector; gag; pol; envelope protein;
KW replication-incompetent retroviral vector; RIR; pCMV-DeltaP;
KW pCMV-DeltaP; pCDNA3-HIVgag-RRE; pCMV-DeltaP;
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX US6451304-B1.
XX
XX 17-SEP-2002.
XX
XX 09-MAR-1999; 99US-0265013.
XX
XX 09-MAR-1998; 98US-077625P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Friedmann T, Miyahara A;
XX
XX WPI; 2003-074128/07.
XX
XX Producing replication-incompetent retrovirus vectors, by transfecting
XX cells with a provirus plasmid that encodes gag, a provirus plasmid that
XX encodes pol, and an envelope protein encoding construct -
XX
XX Example 2; Page -; 16pp; English.
XX
XX The invention relates to producing replication-incompetent retrovirus
XX vectors (RIR), involving transfecting cells with a first provirus plasmid
```

Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

D	b	6939	TCAGCACA	GTACAT	GTACACAT	TGAAT	TGAAT	TGAGG	CCAGTA	AGTA	AT	TCACAT	CAC	TGCT	GT	TTAA	6998
O	y	2071	ATGGCAGT	TCAGC	GAAGAA	GTAGT	AGT	TAGT	TAAT	TTAGAT	CT	TGCCAAT	TTT	CCACAG	CAAT	GT	2130
D	b	6999	ATGGCAGT	CTAGCA	GGAAGAA	GTAGT	AGT	TAAT	TTAGAT	CT	TGCCAAT	TTT	CCACAG	CAAT	GT	2130	
O	y	2131	AAACCAT	TAAT	GTACAG	CTGAC	TGAC	CAAT	CT	GTAGAA	TTAAT	TTAGT	ACAG	ACC	CAAC	CA	2190
D	b	7059	AAACCAT	TAAT	GTACAG	CTGAC	TGAC	CAAT	CT	GTAGAA	TTAAT	TTAGT	ACAG	ACC	CAAC	CA	2190
O	y	2191	AT	NC	AAAA	AAAA	AGAT	CCGTA	CTCAG	AGGGG	ACC	AGGAG	GCAT	TTT	GT	TCAT	2250
D	b	7119	AT	NC	AAAA	AAAA	AGAT	CCGTA	CTCAG	AGGGG	ACC	AGGAG	GCAT	TTT	GT	TCAT	2250
O	y	2251	AAAT	AT	GGA	AAAT	TAT	TAGAC	CAAG	CACT	TT	TAACAT	TAG	AGCA	AAAT	TGCAAT	2310
D	b	7179	AAAT	AT	GGA	AAAT	TAT	TAGAC	CAAG	CACT	TT	TAACAT	TAG	AGCA	AAAT	TGCAAT	2310
O	y	2311	TAAA	AC	GAT	AGT	AGT	AGCA	AAAT	TAA	GAG	AA	CAAT	TTT	GGA	CAAT	2370
D	b	7239	TAAA	AC	GAT	AGT	AGT	AGCA	AAAT	TAA	GAG	AA	CAAT	TTT	GGA	CAAT	2370
O	y	2371	AGCA	AT	CCT	CCAG	AGGG	AGCC	CA	AAAT	TGTA	AG	CGCA	AGT	TTTAAT	TG	2430
D	b	7299	AGCA	AT	CCT	CCAG	AGGG	AGCC	CA	AAAT	TGTA	AG	CGCA	AGT	TTTAAT	TG	2430
O	y	2431	TTT	CT	ACT	GT	AAAT	TC	AAAC	AC	AACT	GT	TTAAT	TAGT	ACT	GT	2490
D	b	7359	TTT	CT	ACT	GT	AAAT	TC	AAAC	AC	AACT	GT	TTAAT	TAGT	ACT	GT	2490
O	y	2491	CTGA	AGG	GT	CA	AAAT	TAC	CTG	AA	GAA	G	TGA	GC	AC	AT	2550
D	b	7419	CTGA	AGG	GT	CA	AAAT	TAC	CTG	AA	GAA	G	TGA	GC	AC	AT	2550
O	y	2551	AAAT	TA	TAA	CA	TGT	GCG	AG	GA	AGT	AT	AG	CC	CT	CC	2610
D	b	7479	AAAT	TA	TAA	CA	TGT	GCG	AG	GA	AGT	AT	AG	CC	CT	CC	2610
O	y	2611	AAAT	T	AGAT	GT	TC	CA	AAAT	TTT	CT	G	GC	T	ATT	TA	2670
D	b	7539	AAAT	T	AGAT	GT	TC	CA	AAAT	TTT	CT	G	GC	T	ATT	TA	2670
O	y	2671	ACA	AT	T	GGG	TC	CG	AGT	CTT	C	AG	CC	T	G	AG	2730
D	b	7599	ACA	AT	T	GGG	TC	CG	AGT	CTT	C	AG	CC	T	G	AG	2730
O	y	2731	AAAT	TA	T	AAAT	TA	TAA	AGT	AGT	AA	AAAT	TG	AA	CA	CTT	2790
D	b	7659	AAAT	TA	T	AAAT	TA	TAA	AGT	AGT	AA	AAAT	TG	AA	CA	CTT	2790
O	y	2791	AG	AG	CA	AA	G	AGT	GGT	CCAG	AAAA	AA	AG	CA	G	AG	2850
D	b	7719	AG	AG	CA	AA	G	AGT	GGT	CCAG	AAAA	AA	AG	CA	G	AG	2850
O	y	2851	TCTT	G	G	G	A	G	C	A	G	A	G	C	A	T	2910
D	b	7779	TCTT	G	G	A	G	C	A	G	A	G	C	A	T	2910	
O	y	2911	GAC	A	T	T	A	T	T	G	T	C	G	A	T	A	2970
D	b	7839	GAC	A	T	T	A	T	T	G	T	C	G	A	T	A	2970
O	y	2971	AAC	G	C	A	T	C	T	G	A	C	T	C	A	C	3030
D	b	7899	AAC	G	C	A	T	C	T	G	A	C	T	C	A	C	3030
O	y	3031	CTG	G	A	A	G	A	T	A	A	G	A	T	A	A	3090
D	b	7959	CTG	G	A												

QY	3151	TTTGGAAATACATGACCTGATGAGTGGACAGAGAAATTACAAATTACAAAGCTTAA	3210	FT	Key	Location/Qualifiers
				FT	repeat_region	1..634
				FT		/*tag= a
				FT		/rpt_type= TERMINAL
				FT		/note= "5'LTR"
QY	3211	TACACTCCTTAATTGAAGATCGCAAAACGACAGAAAGAAATGACAGAAATTATTGG	3270	FT	repeat_unit	456..548
				FT		/*tag= b
				FT		/standard_name= R
				FT		375..385
QY	3271	AATTAGATAAATGGGCAAGTTTGGAATGGTTTAACATTAACAAATGGCTGGTATA	3330	FT	GC_signal	/*tag= c
				FT		/standard_name= Spl_binding_site
				FT		389..395
				FT		/*tag= d
				FT		/standard_name= Spl_binding_site
				FT		399..407
QY	3331	TAAATTAATTCATATGATAGTAGAGGCTTGAGTGTTAAGATAAGTTTTCCTGTAC	3390	FT	GC_signal	/*tag= e
				FT		/standard_name= Spl_binding_site
				FT		636..656
				FT		/*tag= f
				FT		/standard_name= Lys_tRNA_pbs
				FT		790..2292
QY	8259	TAAATTAATTCATATGATAGTAGAGGCTTGAGTGTTAAGATAAGTTTTCCTGTAC	8318	FT	GC_signal	/*tag= g
				FT		/product= gag
				FT		2087..5096
				FT		/*tag= h
				FT		/product= pol
				FT		/note= "NH2-terminal uncertain"
				FT		5041..5619
QY	3451	CAATCCGAGGGGACCGGACGCGCCGAGGAGTAAGAGAGAGTGAGAGAGACA	3510	FT	CDS	/*tag= i
				FT		/product= vif
				FT		5559..5849
				FT		/*tag= j
				FT		/product= vpr
				FT		6061..6306
				FT		/*tag= k
				FT		/product= vpu
				FT		5830..6044
				FT		/*tag= l
				FT		/product= tat
				FT		/note= "full-length tat obtained by splicing"
				FT		5969..6044
				FT		/*tag= m
				FT		/product= rev
				FT		/note= "full-length rev obtained by splicing"
				FT		8369..8414
				FT		/*tag= n
				FT		/product= tat
				FT		/note= "see above"
				FT		8369..8643
				FT		/*tag= o
				FT		/product= rev
				FT		/note= "see above"
				FT		6221..8785
				FT		/*tag= p
				FT		/product= env
				FT		8787..9407
				FT		/*tag= q
				FT		/product= nef
				FT		9076..9709
				FT		/*tag= r
				FT		/rpt_type= TERMINAL
				FT		/note= "3'LTR"
				FT		9531..9624
				FT		/*tag= s
				FT		/standard_name= R
				FT		9602..9607
				FT		/*tag= t
				FT		
RESULT 9						
AAQ22488						
ID	AAQ22488	standard; DNA; 9709 BP.				
XX						
AC	AAQ22488;					
XX						
DT	06-JUL-1992	(first entry)				
XX						
DE	HIV-1 proviral clone pNL4-3.					
XX						
XX						
KM	AIDS; Acquired Immune Deficiency Syndrome; polymerase chain reaction;					
XX	PCR; site-directed mutagenesis; retrovirus; null-mutation; human; ss.					
OS	Human immunodeficiency virus.					
XX						

PH	Key	Location/Qualifiers
FT	repeat_region	1..634
FT		/*tag= a
FT		/rpt_type= TERMINAL
FT		/note= "5'LTR"
FT	repeat_unit	456..548
FT		/*tag= b
FT		/standard_name= R
FT		375..385
FT	GC_signal	/*tag= c
FT		/standard_name= Spl_binding_site
FT		389..395
FT		/*tag= d
FT		/standard_name= Spl_binding_site
FT		399..407
FT	GC_signal	/*tag= e
FT		/standard_name= Spl_binding_site
FT		636..656
FT		/*tag= f
FT		/standard_name= Lys_tRNA_pbs
FT		790..2292
FT	CDS	/*tag= g
FT		/product= gag
FT		2087..5096
FT		/*tag= h
FT		/product= pol
FT		/note= "NH2-terminal uncertain"
FT		5041..5619
FT	CDS	/*tag= i
FT		/product= vif
FT		5559..5849
FT		/*tag= j
FT		/product= vpr
FT		6061..6306
FT	CDS	/*tag= k
FT		/product= vpu
FT		5830..6044
FT		/*tag= l
FT		/product= tat
FT		/note= "full-length tat obtained by splicing"
FT		5969..6044
FT		/*tag= m
FT		/product= rev
FT		/note= "full-length rev obtained by splicing"
FT		8369..8414
FT	exon	/*tag= n
FT		/product= tat
FT		/note= "see above"
FT		8369..8643
FT	exon	/*tag= o
FT		/product= rev
FT		/note= "see above"
FT		6221..8785
FT	CDS	/*tag= p
FT		/product= env
FT		8787..9407
FT		/*tag= q
FT		/product= nef
FT		9076..9709
FT	repeat_region	/*tag= r
FT		/rpt_type= TERMINAL
FT		/note= "3'LTR"
FT		9531..9624
FT	repeat_unit	/*tag= s
FT		/standard_name= R
FT		9602..9607
FT		/*tag= t
PN	polyA_signal	
XX		
XX	MO9200987-A.	
XX	23-JAN-1992.	
PD		
XX	10-JUL-1991;	
PF		91WO-US04884.

XX 12-JUL-1990; 90US-0551945.  
PR (HARD ) HARVARD COLLEGE.  
XX Desrosiers RC;  
XX WPI, 1992-056816/07.  
DR  
XX  
XX Primate lentivirus vaccine protecting against AIDS - and primate  
PT lentiviruses and their DNA clones contg. null mutations, useful for  
PT producing vaccine  
XX  
XX Disclosure; Fig 3; 51pp; English.  
XX  
XX The proviral clone pNL4-3 was used as the basis for creating the  
CC null-mutations of the invention. The clone was described in  
CC Adachi et al., J.Virol. 59:284, 1986. See AA021079-Q21086 for  
CC examples of mutagenic primers for site-directed deletion of regions  
CC of NL4-3.  
XX  
XX Sequence 9709 BP: 3421 A; 1759 C; 2365 G; 2161 T; 3 other:  
SQ  
Query Match 42.9%; Score 2669.8; DB 13; Length 9709;  
Best Local Similarity 99.7%; Pred. No. 3.4e-313;  
Matches 2674; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1291 CCATGAGAGTAAAGAGAGAGATATCAGCCTTGTGAGATGGGGGTGCAATGGGGCACCA 1350  
DB 6219 CAATGAGAGTAAAGAGAGAGATATCAGCCTTGTGAGATGGGGGTGCAATGGGGCACCA 6278  
QY 1351 TGTCTCTGGGATATGATGATCTGTAGTGTCTACAGAAATTTGGGGTCAACCGTCTATT 1410  
DB 6279 TGTCTCTGGGATATGATGATCTGTAGTGTCTACAGAAATTTGGGGTCAACCGTCTATT 6338  
QY 1411 ATGGGGTACTGTGTGGAAGAGCAACACCACTCTATTTTGTGCATCAGATCTTAAAG 1470  
DB 6339 ATGGGGTACTGTGTGGAAGAGCAACACCACTCTATTTTGTGCATCAGATCTTAAAG 6398  
QY 1471 CATATGATACAGAGTACATATGATTTGGGGCACACATGCGTGTACCCACAGACCCCA 1530  
DB 6399 CATATGATACAGAGTACATATGATTTGGGGCACACATGCGTGTACCCACAGACCCCA 6458  
QY 1531 ACCCAAGAGATGATTTGTTAAATGTGACAGAAATTTTAAATGGAAGAAATGACA 1590  
DB 6459 ACCCAAGAGATGATTTGTTAAATGTGACAGAAATTTTAAATGGAAGAAATGACA 6518  
QY 1591 TGTGAGAACAGATCATGAGATATATCACTTTATGGATCAAAAGCCCTAAAGCATGTG 1650  
DB 6519 TGTGAGAACAGATCATGAGATATATCACTTTATGGATCAAAAGCCCTAAAGCATGTG 6578  
QY 1651 TAAATTTAAACCCCACTCTGTGTAGTTTAAAGTGCACATGTTGAAGATATCTATA 1710  
DB 6579 TAAATTTAAACCCCACTCTGTGTAGTTTAAAGTGCACATGTTGAAGATATCTATA 6638  
QY 1711 CCAATAGTAGTAGGGGAGATGATTAATGAGAAAGAGATATAAAACTGCTTCA 1770  
DB 6639 CCAATAGTAGTAGGGGAGATGATTAATGAGAAAGAGATATAAAACTGCTTCA 6698  
QY 1771 ATATCAGCAAGAGATAGAGATTAAGTGCAGAAAGATATGATTTCTTTTAAACTTG 1830  
DB 6699 ATATCAGCAAGAGATAGAGATTAAGTGCAGAAAGATATGATTTCTTTTAAACTTG 6758  
QY 1831 ATATAGTAGCAATATGATATACGCTATAGTGTGATAGTGTGAACACTCAGCATTA 1890  
DB 6759 ATATAGTAGCAATATGATATACGCTATAGTGTGATAGTGTGAACACTCAGCATTA 6818  
QY 1891 CACAGGCTGTCCAAGGATATCTTTGAGCAATTTCCCATCATATTTGGCCGGGTG 1950  
DB 6819 CACAGGCTGTCCAAGGATATCTTTGAGCAATTTCCCATCATATTTGGCCGGGTG 6878  
QY 1951 GTTTTGGCATTTCTAAATGTATATAGAGTTCAATGGAACGAGCATGTCAATG 2010  
|||||

DB 6879 GTTTTGGCATTTCTAAATGTATATATAGAGCTTCAATGGAACAGGACCATGTACAATG 6938  
QY 2011 TCAGCAGCTACATATGATATGAGATGGAATCAGGCCAGTATGATCAACTAATCTCTTTAA 2070  
DB 6939 TCAGCAGCTACATATGATATGAGATGGAATCAGGCCAGTATGATCAACTAATCTCTTTAA 6998  
QY 2071 ATGGCAGCTACGAGAAAGATGATATGATATGATGATGATGATGATGATGATGATGATGAT 2130  
DB 6999 ATGGCAGCTACGAGAAAGATGATATGATATGATGATGATGATGATGATGATGATGATGAT 7058  
QY 2131 AAACCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190  
DB 7059 AAACCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7118  
QY 2191 ATACAGAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2250  
DB 7119 ACACAGAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7178  
QY 2251 AAATAGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2310  
DB 7179 AAATAGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7238  
QY 2311 TAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAATTAATTAATTAATTAATTTTAA 2370  
DB 7239 TAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAATTAATTAATTAATTAATTTTAA 7298  
QY 2371 AGCAATCTCAGAGAGGAGCCAGAAATTTGTAACGACAGTTTAAATTTGGAGGGAT 2430  
DB 7299 AGCAATCTCAGAGAGGAGCCAGAAATTTGTAACGACAGTTTAAATTTGGAGGGAT 7358  
QY 2431 TTTTCTACTGTAATTCACACACACGTTTATAGTGTGTTTAAATGATCTTGGAGTA 2490  
DB 7359 TTTTCTACTGTAATTCACACACACGTTTATAGTGTGTTTAAATGATCTTGGAGTA 7418  
QY 2491 CTGAAGGCTCAATPACACTGAGAGAAAGTGAACATCAGCTCCATGAGAAATTAAC 2550  
DB 7419 CTGAAGGCTCAATPACACTGAGAGAAAGTGAACATCAGCTCCATGAGAAATTAAC 7478  
QY 2551 AATTATTAACATGAGGAGAGTGAAGAAAGCAATGTAATGAGGAGGAGGAGGAGGAGGAGGAG 2610  
DB 7479 AATTATTAACATGAGGAGAGTGAAGAAAGCAATGTAATGAGGAGGAGGAGGAGGAGGAGGAG 7538  
QY 2611 AATTATTAACATGAGGAGAGTGAAGAAAGCAATGTAATGAGGAGGAGGAGGAGGAGGAGGAG 2670  
DB 7539 AATTATTAACATGAGGAGAGTGAAGAAAGCAATGTAATGAGGAGGAGGAGGAGGAGGAGGAG 7598  
QY 2671 ACAATGGCTCGAGATCTTCAAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2730  
DB 7599 ACAATGGCTCGAGATCTTCAAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7658  
QY 2731 AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2790  
DB 7659 AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7718  
QY 2791 AGAGAAGAGTGTGAG 2850  
DB 7719 AGAGAAGAGTGTGAG 7778  
QY 2851 TCTTGGAG 2910  
DB 7779 TCTTGGAG 7838  
QY 2911 GACAATTTATTTGTCTGATATAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2970  
DB 7839 GACAATTTATTTGTCTGATATAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7898  
QY 2971 AACAGCATCTGTTCACTCAGCTCAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3030  
DB 7899 AACAGCATCTGTTCACTCAGCTCAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7958  
QY 3031 CTGTGAAAGATATACCTAAAGAGATCAACAGCTCCTGGGGAGTTTGGGGTGTCTGTGAAAG 3090  
DB 7959 CTGTGAAAGATATACCTAAAGAGATCAACAGCTCCTGGGGAGTTTGGGGTGTCTGTGAAAG 8018  
|||||





|||||  
Db 6644 TGGTGAACAGATGATGAGATATAATCAGTTTATGGGATCAAGCCATAAGCCATGTG 6703  
QY 1651 TAAATTTAACCCACTCTGTGTGTACTTTAAAGTCAGTGTATTGAAGATGATCTATA 1710  
Db 6704 TAAATTTAACCCACTCTGTGTGTATTAAAGTCAGTGTATTGAAGATGATCTATA 6763  
QY 1711 CCAATAGTAGTGGGGGAATGATTAATGGAGAAAGAGATTAATAAAAGCCCTCTTCA 1770  
Db 6764 CCAATAGTAGTGGGGGAATGATTAATGGAGAAAGAGATTAATAAAAGCCCTCTTCA 6823  
QY 1771 ATATCAGCAAGCAAGTAAAGATTAAGTGCAGAAAGATATGATCTTTTATAACTTG 1830  
Db 6824 ATATCAGCAAGCAAGTAAAGATTAAGTGCAGAAAGATATGATCTTTTATAACTTG 6883  
QY 1831 ATATAGTACCAATAGATATAATACCAAGCTATAGTGTATAGTTGATAACCTCAGTCATTA 1890  
Db 6884 ATATAGTACCAATAGATATAATACCAAGCTATAGTGTATAGTTGATAACCTCAGTCATTA 6943  
QY 1891 CACAGGCTGTCCAAAGGTATCCCTTGAGCCCAATCCCATACATTAATGTGCCCCGGCTG 1950  
Db 6944 CACAGGCTGTCCAAAGGTATCCCTTGAGCCCAATCCCATACATTAATGTGCCCCGGCTG 7003  
QY 1951 GTTTTGCATTTCTAAATGTATATAATTAAGAGCTTCAATGGAACAGAGCCATGTCAAA 2010  
Db 7004 GTTTTGCATTTCTAAATGTATATAATTAAGAGCTTCAATGGAACAGAGCCATGTCAAA 7063  
QY 2011 TCAGACAGATCAATGTATACATGTGAATCAGGCCAGTGTATCAACTCACTGCTGTAA 2070  
Db 7064 TCAGACAGATCAATGTATACATGTGAATCAGGCCAGTGTATCAACTCACTGCTGTAA 7123  
QY 2071 ATGGAGTCTAGCAGAAAGATGTAGTAAATAGATGCCCAATTTCCAGACATGTGA 2130  
Db 7124 ATGGAGTCTAGCAGAAAGATGTAGTAAATAGATGCCCAATTTCCAGACATGTGA 7183  
QY 2131 AAACCATATAGTACAGCTGAACACATCTGTGAATTAATTTGACAAGACCCAAACA 2190  
Db 7184 AAACCATATAGTACAGCTGAACACATCTGTGAATTAATTTGACAAGACCCAAACA 7243  
QY 2191 ATACAAGAAAAAGTATCCGTATCCAGAGGGACCCAGGAGACATTTGTTACATAGAA 2250  
Db 7244 ATACAAGAAAAAGTATCCGTATCCAGAGGGACCCAGGAGACATTTGTTACATAGAA 7303  
QY 2251 AAATGGAATGTAGACAAGCAATGTGAACATTAGTAGACAAATGGATGCACTT 2310  
Db 7304 AAATGGAATGTAGACAAGCAATGTGAACATTAGTAGAGCAAAATGGATGCACTT 7363  
QY 2311 TAAACAGATAGCTAGCAAAATTAAGAAACAATTTGAAATATATAAACAATATCTTTA 2370  
Db 7364 TAAACAGATAGCTAGCAAAATTAAGAAACAATTTGAAATATATAAACAATATCTTTA 7423  
QY 2371 AGCAATCTCAGAGAGGGACCCAGAAATTTGAAGCAGCATTTTAAATGTGGAGGGAA 2430  
Db 7424 AGCAATCTCAGAGAGGGACCCAGAAATTTGAAGCAGCATTTTAAATGTGGAGGGAA 7483  
QY 2431 TTTTCTAGCTAATTCACACACACAGTGTAAATAGTACTGTTTAAATAGTACTTGGAGTA 2490  
Db 7484 TTTTCTAGCTAATTCACACACACAGTGTAAATAGTACTGTTTAAATAGTACTTGGAGTA 7543  
QY 2491 CTGAAGGGTCAATTAACACTGAAGAAAGTGAACATCACTCCCATGCAAAATAAAC 2550  
Db 7544 CTGAAGGGTCAATTAACACTGAAGAAAGTGAACATCACTCCCATGCAAAATAAAC 7603  
QY 2551 AATTTATTAACATGTGGCAGAGATGAAGAAAAAGCAATGTATGCCCTCCCATAGTGAC 2610  
Db 7604 AATTTATTAACATGTGGCAGAGATGAAGAAAAAGCAATGTATGCCCTCCCATAGTGAC 7663  
QY 2611 AATTAGATGTCAATCAATATATTACTGGGCTGCTTTTAAAGAGATGTTGTAATPACA 2670  
Db 7664 AATTAGATGTCAATCAATATATTACTGGGCTGCTTTTAAAGAGATGTTGTAATPACA 7723  
QY 2671 ACAATGGGTCCGAGATCTTCAGACCTGAGAGGGGATGTGAGGACAAATTTGAGAAAGTG 2730  
|||||

|||||  
Db 7724 ACAATGGGTCCGAGATCTTCAGACCTGAGAGAGCCGATATGAGGACAAATTTGAGAAAGTG 7783  
QY 2731 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2790  
Db 7784 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7843  
QY 2791 AGAGAAAGTGGTGCAGAGAGAAAAAGAGAGTGGAAATAGAGACTTTGTTCCCTGGGT 2850  
Db 7844 AGAGAAAGTGGTGCAGAGAGAAAAAGAGAGTGGAAATAGAGACTTTGTTCCCTGGGT 7903  
QY 2851 TCTTGGAGCAGCAGAGAAAGCACTATGGGCTGCAGCTCAATGACGCTGAGCGGTACAGCCA 2910  
Db 7904 TCTTGGAGCAGCAGAGAAAGCACTATGGGCTGCAGCTCAATGACGCTGAGCGGTACAGCCA 7963  
QY 2911 GACATTTATTTGCTGTATATAGTGCAGCAGCAGAACAAATTTGCTGAGGGCTATTGAGGGCC 2970  
Db 7964 GACATTTATTTGCTGTATATAGTGCAGCAGCAGAACAAATTTGCTGAGGGCTATTGAGGGCC 8023  
QY 2971 AACAGCATCTGTGCACTCAGAGTCTGGGGCATCAACAGCTCCAGGCCAAGATCTGCG 3030  
Db 8024 AACAGCATCTGTGCACTCAGAGTCTGGGGCATCAACAGCTCCAGGCCAAGATCTGCG 8083  
QY 3031 CTGTGAAAAGTACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTGGAANAAC 3090  
Db 8084 CTGTGAAAAGTACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTGGAANAAC 8143  
QY 3091 TCATTTGACACACCTGCTGCTGGGATGAGTGTAGTAAATCTGGAANAAC 3150  
Db 8144 TCATTTGACACACCTGCTGCTGGGATGAGTGTAGTAAATCTGGAANAAC 8203  
QY 3151 TTTGGAATTAACATGACCTGATGAGTGGAGAGAAATTAACAAATTAACAACTTAA 3210  
Db 8204 TTTGGAATTAACATGACCTGATGAGTGGAGAGAAATTAACAAATTAACAACTTAA 8263  
QY 3211 TACACTCTTAATTAAGAAATCGCAAAACCCAGAAAGAAATGATACAAATTAATTTGG 3270  
Db 8264 TACACTCTTAATTAAGAAATCGCAAAACCCAGAAAGAAATGATACAAATTAATTTGG 8323  
QY 3271 AATTAAGTAAATGGGCAAGTTGTGGAATTTGTTAACATTAACAAATTTGGCTGTATA 3330  
Db 8324 AATTAAGTAAATGGGCAAGTTGTGGAATTTGTTAACATTAACAAATTTGGCTGTATA 8383  
QY 3331 TAAATTTATTCATTAATGATAGTAGAGAGCTTGTGAGTTTAAATAGTTTTCGCTGAC 3390  
Db 8384 TAAATTTATTCATTAATGATAGTAGAGAGCTTGTGAGTTTAAATAGTTTTCGCTGAC 8443  
QY 3391 TTTCTATAGTGAATAGAGTTAGGAGGATATTCACCATTAATGTTTGAAGCCACTCC 3450  
Db 8444 TTTCTATAGTGAATAGAGTTAGGAGGATATTCACCATTAATGTTTGAAGCCACTCC 8503  
QY 3451 CAATCCCGAGGGGACCCGACAGGCCGAAAGAAATGAAGAAAGAGTGGAGAGAGACA 3510  
Db 8504 CAATCCCGAGGGGACCCGACAGGCCGAAAGAAATGAAGAAAGAGTGGAGAGAGACA 8563  
QY 3511 GAGACAGATCCATTTGATAGTGAACGAGTCTTGAACCTTTATCTGGAGCATCTGCGGA 3570  
Db 8564 GAGACAGATCCATTTGATAGTGAACGAGTCTTGAACCTTTATCTGGAGCATCTGCGGA 8623  
QY 3571 GCCCTGTGCTTTCAGCTACACCGCTTGAGAGACTTACTCTTGATTTGAAGAGAGATTG 3630  
Db 8624 GCCCTGTGCTTTCAGCTACACCGCTTGAGAGACTTACTCTTGATTTGAAGAGAGATTG 8683  
QY 3631 TGGAACTTCTGGGACGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCTACAT 3690  
Db 8684 TGGAACTTCTGGGACGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCTACAT 8743  
QY 3691 ATTGAGTCAGAGAACTAAAGATATGCTGTTAACTTGCATATCCAGCAGCATAGAG 3750  
Db 8744 ATTGAGTCAGAGAACTAAAGATATGCTGTTAACTTGCATATCCAGCAGCATAGAG 8803  
QY 3751 TAGCTAGGAGGACAGATAGGGTTATAGAATTAACAAGCACTTATAGAGCTATTGCGC 3810  
Db 8804 TAGCTAGGAGGACAGATAGGGTTATAGAATTAACAAGCACTTATAGAGCTATTGCGC 8863  
|||||

QY	3811	ACATACCTAGAGAAGATATAGACAGGGCTTGGAAAGGATTTTCTCTAATGATGGGTGGCAAG	3870
DB	8864	ACATACCTAGAGAAGATATAGACAGGGCTTGGAAAGGATTTTCTCTAATGATGGGTGGCAAG	89233
QY	3871	TGTCGTAAGAAATAGTGTGATTTGATGGGCTCTCTTAAGAGGAAAGAGAGAGCTGAG	3930
DB	8924	TGTCGTAAGAAATAGTGTGATTTGATGGGCTCTCTTAAGAGGAAAGAGAGAGCTGAG	8983
QY	3931	CCAGCAGCAGATGGGGTGGAGCGACGATCTCTGAGATCTAGA	3971
DB	8984	CCAGCAGCAGATGGGGTGGAGCGACGATCTCTGAGATCTAGA	9024
RESULT 11			
AAV14355			
ID	AAV14355	standard; DNA; 15581 BP.	
XX	AAV14355;		
AC	19-MAY-1998	(first entry)	
DT			
XX			
DE		Plasmid pMLnSG11 expressing green fluorescent protein.	
XX			
KW		Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish;	
KM		genetic engineering marker; gene therapy; plasmid; ss.	
OS		Synthetic.	
XX			
PN	MO9742320-A1.		
PD	13-NOV-1997.		
XX			
PF	07-MAY-1997;	97MO-USO7625.	
PR	08-MAY-1996;	96US-0646538.	
XX			
PA	(USSH ) US SEC DEPT HEALTH.		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
XX			
PI	Galtanaris GA, Pavlakis GN, Stauber RH, Vournakis JN;		
DR	WPI; 1997-558982/51.		
XX			
PT	New nucleic acid encoding proteins of Aequorea victoria with		
PT	increased fluorescence - useful as markers for detecting cellular		
PT	transformation, subcellular localisation of proteins, for assessing		
PT	gene therapy, mutagenicity etc.		
XX			
PS	Example 13; Page 89-96; 105pp; English.		
XX			
CC	This sequence is a plasmid used to express the mutated green fluorescent		
CC	protein (GFP) of Aequorea victoria coding sequence of the invention. The		
CC	DNA of the invention that encodes a protein that is (a) a protein that		
CC	has leu at position 65 (and optionally Thr at 168 and optionally further		
CC	Cys at 66) and has cellular fluorescence at least 5 times that of		
CC	wild-type GFP or (b) is a blue fluorescent protein (BFP) with His at		
CC	position 67 and also at least one of Leu at 65 and Ala at 164, and has		
CC	cellular fluorescence at least 5 times that of BFP (Tyr67 to His). The		
CC	nucleic acids can be used as markers in genetic engineering and gene		
CC	therapy. They may also be used to detect and characterise regulatory and		
CC	coding sequence elements that control subcellular expression and		
CC	targeting of proteins. Typical applications are monitoring targeting and		
CC	transport of proteins in cells; assessment of gene therapy procedures; in		
CC	diagnosis (when expressed under control of a promoter induced by a		
CC	particular analyte); assessment of mutagenicity of compounds; and for		
CC	drug screening (where expression is controlled by the promoter of a		
CC	target gene), particularly for antiviral or antiparasitic agents.		
XX			
SO	Sequence 15581 BP; 4975 A; 3196 C; 3767 G; 3643 T; 0 other;		
Query Match			
42.3%; Score 2636.6; DB 18; Length 15581;			
Best Local Similarity 99.3%; Pred. No. 2.9e-309;			

OY 2371 AGCAATCCCTCAGAGGGAGCCAGAAATTTGAACGACAGTTTATTTGAGGGGAAT 2430  
 |||||  
 Db 7299 AGCATCCTCAGAGAGGAGCCAGAAATTTGAACGACAGTTTATTTGAGGGGAAT 7358  
 OY 2431 TTTTCTACTGTAAATTCACACAACTGTTTAATAGTACTGTTTAATAGTACTGAGTA 2490  
 |||||  
 Db 7359 TTTTCTACTGTAAATTCACACAACTGTTTAATAGTACTGTTTAATAGTACTGAGTA 7418  
 OY 2491 CTGAAGGTCMAATTAACACTGAAGAAATGACACAACTCCCATGACAGATTAAC 2550  
 |||||  
 Db 7419 CTGAAGGTCMAATTAACACTGAAGAAATGACACAACTCCCATGACAGATTAAC 7478  
 OY 2551 AATTTATTAACATGTCGAGAGAGTGAAGAAAAGCAATGTATGCCCTCCCATGAGTGC 2610  
 |||||  
 Db 7479 AATTTATTAACATGTCGAGAGAGTGAAGAAAAGCAATGTATGCCCTCCCATGAGTGC 7538  
 OY 2611 AATTTAGATGTTTCATCAAAATTTACTGGGCTGCTATTAACAAGAGATGGTGAATACA 2670  
 |||||  
 Db 7539 AATTTAGATGTTTCATCAAAATTTACTGGGCTGCTATTAACAAGAGATGGTGAATACA 7598  
 OY 2671 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGGAGATATGAGGACAAATTTGAGAAATG 2730  
 |||||  
 Db 7599 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGGAGATATGAGGACAAATTTGAGAAATG 7658  
 OY 2731 AATTTATTAATATTAATTAAGTAAATTTGAACCATTTAGAGTAGCACCCCAAGGCA 2790  
 |||||  
 Db 7659 AATTTATTAATATTAATTAAGTAAATTTGAACCATTTAGAGTAGCACCCCAAGGCA 7718  
 OY 2791 AGAGAAGATGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTGTTCCTTGGGT 2850  
 |||||  
 Db 7719 AGAGAAGATGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTGTTCCTTGGGT 7778  
 OY 2851 TCTTGGGAGCAGAGAGAAAGCAGTATGGGCTGCACGTCATAGCCTGACGGTACAGGCCA 2910  
 |||||  
 Db 7779 TCTTGGGAGCAGAGAGAAAGCAGTATGGGCTGCACGTCATAGCCTGACGGTACAGGCCA 7838  
 OY 2911 GACATTTATTTGCTGATATAGTGCAGAGCAAGAACATTTGCTGAGGGCTATTTGAGGGCC 2970  
 |||||  
 Db 7839 GACATTTATTTGCTGATATAGTGCAGAGCAAGAACATTTGCTGAGGGCTATTTGAGGGCC 7898  
 OY 2971 AACAGCATCTGTGGCACTCAACAGTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 3030  
 |||||  
 Db 7899 AACAGCATCTGTGGCACTCAACAGTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 7958  
 OY 3031 CTGTGGAAGATACCTTAAGAGATCAACAGCTCTGTGGGATTTGGGCTTCTGGAAGC 3090  
 |||||  
 Db 7959 CTGTGGAAGATACCTTAAGAGATCAACAGCTCTGTGGGATTTGGGCTTCTGGAAGC 8018  
 OY 3091 TCATTTGCACCACTGCTGTGCTGGATGCTAGTTGAGTAATTAATCTCTGGAAGCAGA 3150  
 |||||  
 Db 8019 TCATTTGCACCACTGCTGTGCTGGATGCTAGTTGAGTAATTAATCTCTGGAAGCAGA 8078  
 OY 3151 TTTTGAATTAACATGACCTGATGAGTGGAGCAGAGAAATTAACAAATTAACAAGCTTAA 3210  
 |||||  
 Db 8079 TTTTGAATTAACATGACCTGATGAGTGGAGCAGAGAAATTAACAATTAACAAGCTTAA 8138  
 OY 3211 TACACTCCTTAATTAAGAAATTCGCAAAACGAGCAAGAAAGATGACAGAAATTTATGG 3270  
 |||||  
 Db 8139 TACACTCCTTAATTAAGAAATTCGCAAAACGAGCAAGAAAGATGACAGAAATTTATGG 8198  
 OY 3271 AATTTAGATTAATGGCAAGTTTGTGAGTATGTTTAACATTAACAATTTGGCTGGGTATA 3330  
 |||||  
 Db 8199 AATTTAGATTAATGGCAAGTTTGTGAGTATGTTTAACATTAACAATTTGGCTGGGTATA 8258  
 OY 3331 TAAATTAATTAATTAAGTATGAGAGGCTTGTAGTATTAAGATTAAGTATTTTGTGTAC 3390  
 |||||  
 Db 8259 TAAATTAATTAATTAAGTATGAGAGGCTTGTAGTATTAAGATTAAGTATTTTGTGTAC 8318  
 OY 3391 TTTTCAATGATTAAGATTAAGGCAAGGATATTAACCTTATCGTTTACAGCCACCTCC 3450  
 |||||  
 Db 8319 TTTTCAATGATTAAGATTAAGGCAAGGATATTAACCTTATCGTTTACAGCCACCTCC 8378

OY 3451 CAATCCGAGGGAGCCGACAGGCCCCGAGAGATAGAAAGAAAGTGGAGAGAGACA 3510  
 |||||  
 Db 8379 CAATCCGAGGGAGCCGACAGGCCCCGAGAGATAGAAAGAAAGTGGAGAGAGACA 8438  
 OY 3511 GAGACAGATTCATTCGATTTAGTGAAGGATCTTAAGCACTTATCTGGAGATTCGCGA 3570  
 |||||  
 Db 8439 GAGACAGATTCATTCGATTTAGTGAAGGATCTTAAGCACTTATCTGGAGATTCGCGA 8498  
 OY 3571 GCCTGTGCTCTTCACGCTACCAACCCCTTGAGAGACTTACTCTTGATTTGAAGAGATTG 3630  
 |||||  
 Db 8499 GCCTGTGCTCTTCACGCTACCAACCCCTTGAGAGACTTACTCTTGATTTGAAGAGATTG 8558  
 OY 3631 TGGAACTCTTGGAGAGCGAGGGGTGGGAAGCCCTCAATATTTGGGAATCTCCACAGT 3690  
 |||||  
 Db 8559 TGGAACTCTTGGAGAGCGAGGGGTGGGAAGCCCTCAATATTTGGGAATCTCCACAGT 8618  
 OY 3691 ATTGGAGTCAGAACTAAAGAAATAGTGTCTTTAACTTCTCAATGCCACAGCCATAGCAG 3750  
 |||||  
 Db 8619 ATTGGAGTCAGAACTAAAGAAATAGTGTCTTTAACTTCTCAATGCCACAGCCATAGCAG 8678  
 OY 3751 TAGCTGAGGGAGCAGATAGGCTTATAGAAATTTACAGACGCTTATAGACTATTCGCC 3810  
 |||||  
 Db 8679 TAGCTGAGGGAGCAGATAGGCTTATAGAAATTTACAGACGCTTATAGACTATTCGCC 8738  
 OY 3811 ACATACCTAGAAAGATTAAGACAGAGGCTTGGAAAGGATTTTGCTATTAAGATGGGTGCAAG 3870  
 |||||  
 Db 8739 ACATACCTAGAAAGATTAAGACAGAGGCTTGGAAAGGATTTTGCTATTAAGATGGGTGCAAG 8798  
 OY 3871 TGGTCAAAAAGTAGTGTATTTGATGGCTGCTGTAAAGGAAAGAAATGAGACGACTGAG 3930  
 |||||  
 Db 8799 TGGTCAAAAAGTAGTGTATTTGATGGCTGCTGTAAAGGAAAGAAATGAGACGACTGAG 8858  
 OY 3931 CCAGCAGCAGATGGGGTGGAGCACTA 3957  
 |||||  
 Db 8859 CAAGAAATGGCTAGCAAGAGAGAAAGAA 8885  
 |||||

RESULT 12  
 AAV74271  
 ID AAV74271 standard; DNA; 9213 BP.  
 XX  
 AC AAV74271;  
 XX  
 DT 14-JUN-1999 (first entry)  
 XX  
 DE HTLV-III genomic DNA.  
 XX  
 KW AIDS; acquired immunodeficiency syndrome; viral infection; env protein;  
 KW fusion polypeptide; E' protein; gag protein; pol protein; p' protein;  
 KW gp signal peptide; detection; vaccination; etiological agent;  
 KW infection inhibitor; AIDS associated retrovirus; p24; gp41; ss.  
 XX  
 OS Human lymphotropic virus type III.  
 XX  
 FH Key  
 FT CDS location/Qualifiers  
 FT  
 FT /\*tag= a  
 FT /product= "gag protein"  
 FT 732..1772  
 FT /\*tag= b  
 FT /product= "p24 gag protein"  
 FT 1639..4677  
 FT /\*tag= c  
 FT /product= "pol protein"  
 FT 4622..5200  
 FT /\*tag= d  
 FT /product= "p' protein"  
 FT 5803..8373  
 FT /\*tag= e  
 FT /product= "env protein"  
 FT 7336..8373  
 FT /\*tag= f  
 FT /product= "gp41 env protein"



```

Db      7181 ATARCAACATGAGTCGAGATCTTCACAGCCTGGAGGAGAGATATGAGGACCAATTGGA 7240
Oy      2725 GAATGGAATTATATTAATATATAAGTAGTAAAAATTGAACCTTAGAGTAGACCCACCA 2784
Db      7241 GAATGGAATTATATTAATATATAAGTAGTAAAAATTGAACCTTAGAGTAGACCCACCA 7300
Oy      2785 AGGCAGAGAGAGAGAGTGTGAGAGAGAAAAAGAGCGTGGGATATGAGCTTTGTTCC 2844
Db      7301 AGGCAGAGAGAGAGTGTGAGAGAGAAAAAGAGCGTGGGATATGAGCTTTGTTCC 7360
Oy      2845 TTGGGTTCTTGGAGCAGCAGCAGAGCAGTATGGCTGCACGTCAATGACGCTGACGTTAC 2904
Db      7361 TTGGGTTCTTGGAGCAGCAGCAGAGCAGTATGGGCGCAGCGTCATGACGCTGACGTTAC 7420
Oy      2905 AGGCCAGACAAATTATTTCTATATAGTGCAGCAGCAGACAAATTTGCTGAGGGCTATTTG 2964
Db      7421 AGGCCAGACAAATTATTTCTGATATAGTGCAGCAGCAGACAAATTTGCTGAGGGCTATTTG 7480
Oy      2965 AGGCCAGACACATCTGTTCAGCTCAGAGTCTGGGGGCTCAAAAGCTCCAGGCAAGAA 3024
Db      7481 AGGCCAGACACATCTGTTCAGCTCAGAGTCTGGGGGCTCAAAAGCTCCAGGCAAGAA 7540
Oy      3025 TCCTGGCTGTGAAAAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGGTTGCTCTG 3084
Db      7541 TCCTGGCTGTGAAAAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGGTTGCTCTG 7600
Oy      3085 GAAACATCATTTGCACCACTCTGTGCTTGAATGCTAGTTGGAGTATATTAATCTCTGG 3144
Db      7601 GAAACATCATTTGCACCACTCTGTGCTTGAATGCTAGTTGGAGTATATTAATCTCTGG 7660
Oy      3145 AACAGATTGGATATACATGACATGATGATGAGTGGGAGAGAGAAATTAACAATTACACAA 3204
Db      7661 AACAGATTGGATATACATGACATGATGATGAGTGGGAGAGAGAAATTAACAATTACACAA 7720
Oy      3205 GCTTAATACACTCTCTTAATTAAGATTCGCAAAACCAAGCAAGAAAGATGACAGAGAT 3264
Db      7721 GCTTAATACACTCTCTTAATTAAGATTCGCAAAACCAAGCAAGAAAGATGACAGAGAT 7780
Oy      3265 TATTGGAATTAGATTAATAGGCAAGTTGTGGAATTTGTTAACATAACAATTGGCTGT 3324
Db      7781 TATTGGAATTAGATTAATAGGCAAGTTGTGGAATTTGTTAACATAACAATTGGCTGT 7840
Oy      3325 GGTATATTAATTAATTAATTAATAGTAGAGAGGCTTGGTAAAGATAGTTTGG 3384
Db      7841 GGTATATTAATTAATTAATTAATAGTAGAGAGGCTTGGTAAAGATAGTTTGG 7900
Oy      3385 CTGACTTTCTATAGTAGATAGATAGAGTAGGAGGATATTCACCAATTATCGTTTGACACC 3444
Db      7901 CTGACTTTCTATAGTAGATAGATAGAGTAGGAGGATATTCACCAATTATCGTTTGACACC 7960
Oy      3445 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCCGAAGGAATGAAGAAAGATGAGAGAGA 3504
Db      7961 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCCGAAGGAATGAAGAAAGATGAGAGAGA 8020
Oy      3505 GAGACAGAGACAGATCCATTCGATTAGTGAACGAGTCTTTCGACCTTCTGGGAGCATTC 3564
Db      8021 GAGACAGAGACAGATCCATTCGATTAGTGAACGAGTCTTTCGACCTTCTGGGAGCATTC 8080
Oy      3565 TGGCGAGCTGTGCTCTTCAGTACACCGCTTGAGAGACTTACTTTGATTGTAAACGA 3624
Db      8081 TGGCGAGCTGTGCTCTTCAGTACACCGCTTGAGAGACTTACTTTGATTGTAAACGA 8140
Oy      3625 GGATTGTGGAATCTTGGGAGCAGAGGGGTGGGAAAGCCCTCAATATTTGGTGAATCTCC 3684
Db      8141 GGATTGTGGAATCTTGGGAGCAGAGGGGTGGGAAAGCCCTCAATATTTGGTGAATCTCC 8200
Oy      3685 TACAGTATTTGAGTCAGAGAACTAAAGAAATAGTGCCTTAATCTGCATATCCACAGCA 3744
Db      8201 TACAGTATTTGAGTCAGAGAACTAAAGAAATAGTGCCTTAATCTGCATATCCACAGCA 8260
Oy      3745 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTTACAGCAGCTTATAGAGCTA 3804
Db      8261 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTTACAGCAGCTTATAGAGCTA 8320

```

```

Oy      3805 TTGGCCACATACCTAGAGAAATAGACACAGGCTTGGAAAGATTTTGTATAGATGGCT 3864
Db      8321 TTGGCCACATACCTAGAGAAATAGACACAGGCTTGGAAAGATTTTGTATAGATGGCT 8380
Oy      3865 GGCAAGTGTGTAAAAAGATGATGATTTGATGGCCCTGCTGAAGGAAAGATGAGACGA 3924
Db      8381 GGCAAGTGTGTAAAAAGATGATGATTTGATGGCCCTGCTGAAGGAAAGATGAGACGA 8440
Oy      3925 GCTGAGCAGCAGCAGAGATGGGTGGGAGCAGTATCTCGAGATCTAGA 3971
Db      8441 GCTGAGCAGCAGCAGAGATGGGTGGGAGCAGTATCTCGAGACCTAGA 8487

RESULT 13
AAV81866
ID      AAV81866 standard; DNA; 9213 BP.
XX
AC      AAV81866;
XX
DT      01-JUN-1999 (first entry)
XX
DE      Nucleic acid sequence of an HIV-1 isolate.
XX
KW      Antigenic composition; primate; lentivirus; nef gene; vaccine;
KW      infection; AIDS; HIV-1; ss.
XX
OS      Human immunodeficiency virus type 1.
XX
PN      US851813-A.
XX
PD      22-DEC-1998.
XX
PE      27-JAN-1994; 94US-0188583.
XX
PR      27-JAN-1994; 94US-0188583.
PR      12-JUL-1994; 90US-0551945.
PR      09-JUL-1991; 91US-0727494.
XX
PA      (HARD ) HARVARD COLLEGE.
XX
PI      Destroiers RC;
XX
DR      WPI: 1999-080408/07.
XX
P      P-PSDB: AAM89322; AAM89323; AAM89324; AAM89325; AAM89326.
XX
PT      Lentivirus antigenic compositions - containing lentivirus with nef
PT      gene deletion
XX
PS      Disclosure: Fig 2A-R; 93pp; English.
XX
CC      The invention relates to an antigenic composition comprising an isolated
CC      primate lentivirus whose genome contains an engineered non-revertible
CC      null mutation in the nef gene, or an infectious DNA clone in a carrier.
CC      The antigenic composition is used in vaccines against infection by the
CC      lentivirus, e.g. AIDS. The present sequence represents the nucleic
CC      acid sequence of an HIV-1 isolate.
XX
SQ      Sequence 9213 BP; 3297 A; 1655 C; 2218 G; 2043 T; 0 other;
XX

Query Match      41.7%; Score 2599.4; DB 20; Length 9213;
Best Local Similarity 98.3%; Pred. No. 1e-304;
Matches 2640; Conservative 0; Mismatches 41; Indels 6; Gaps 1;

Oy      1291 CCATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGTGGAAATGGGGCACCA 1350
Db      5801 CAATGAGAGTGAAGAGAGAAATATCAGCACTTGTGAGATGGGGTGGAGATGGGGCACCA 5860
Oy      1351 TGTCTCTTGGGATATGATGATGATCTGTAGCTTACAGAAAAATTTGGGCTCACGCTATTT 1410
Db      5861 TGTCTCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5920
Oy      1411 ATGGGGTACCTGTGTGGAAGAACCAACCACTCATTTTGTGATCATCATGCTAAG 1470

```

Db 5921 ATGGGGTACCTGTGTGGAAGAGCAACACCACTATTTTGTGATCGAGATGTAAG 5980  
QY 1471 CATATGATACAGAGGTGATCAATATGTTTGGGCCACACATGCTGTGTAACCCACAGACCCA 1530  
Db 5981 CATATATACAGAGGTGATCAATATGTTTGGGCCACACATGCTGTGTAACCCACAGACCCA 6040  
QY 1531 ACCCACAAGAAAGTATGTTGTTAAATGTGCACAGAAATTTTAAATCATGTGGAAATAGCA 1590  
Db 6041 ACCCACAAGAGTATGTTGTTAAATGTGCACAGAAATTTTAAATCATGTGGAAATAGCA 6100  
QY 1591 TGGTAGACAGATGCATGAGGATATATCATGTTTATGGATCAAAAGCCTAAAGCCATGTG 1650  
Db 6101 TGGTAGACAGATGCATGAGGATATATCATGTTTATGGATCAAAAGCCTAAAGCCATGTG 6160  
QY 1651 TAAATTTAACCCCACTCTGTGTGTTAAAGTGCACTGATTTGAAGAAATGATCTATATA 1710  
Db 6161 TAAATTTAACCCCACTCTGTGTGTTAAAGTGCACTGATTTGAAGAAATGATCTATATA 6220  
QY 1711 CCAATAGTAGTAGCGGAGAAATGATATATGAGAAAGAGATAAAAAACTGCTTTTCA 1770  
Db 6221 CCAATAGTAGTAGCGGAGAAATGATATATGAGAAAGAGATAAAAAACTGCTTTTCA 6280  
QY 1771 ATATCAGCACAGCATATAGAGATTAAGGTGCAAGAAAGATATGCAATCTTTTATTAACCTTG 1830  
Db 6281 ATATCAGCACAGCATATAGAGATTAAGGTGCAAGAAAGATATGCAATCTTTTATTAACCTTG 6340  
QY 1831 ATATAGTACCAATAGATATA-----TACCACTATAGTGTGATAGTGTGAACACCTCAG 1884  
Db 6341 ATATATATCCAAATAGATATAGATATCTACAGCTATACCTGTGACAACTGTGAACACCTCAG 6400  
QY 1885 TCATTTACACAGGCGCTGCCAAAGGTATCCTTTGAGCCAAATTCCTATCATTTATTTGCCCC 1944  
Db 6401 TCATTTACACAGGCGCTGCCAAAGGTATCCTTTGAGCCAAATTCCTATCATTTATTTGCCCC 6460  
QY 1945 CGGCTGGTTTTGGGATTTCTAAAAATGTAATATTAAGACGTTCAATGGAACAGACATGTA 2004  
Db 6461 CGGCTGGTTTTGGGATTTCTAAAAATGTAATATTAAGACGTTCAATGGAACAGACATGTA 6520  
QY 2005 CAATATGCACAGTACATGTAATGTAACATGGAATCAGGCGAGTATGTTTCAACTCAATGTC 2064  
Db 6521 CAATATGCACAGTACATGTAATGTAACATGGAATGAGCCAGTATGTTTCAACTCAATGTC 6580  
QY 2065 TGTAAATGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAATTTTACAGACA 2124  
Db 6581 TGTAAATGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAATTTTACAGACA 6640  
QY 2125 ATGCTTAAACCATATATGTAACAGCTGAAACATCTGTGAATTTAATTTGTAACAGACCA 2184  
Db 6641 ATGCTTAAACCATATATGTAACAGCTGAAACATCTGTGAATTTAATTTGTAACAGACCA 6700  
QY 2185 ACAACATATCAAGAAAAAGTATCGTATCCAGAGGGGACCAAGGAGGCAATTTGTTTCAA 2244  
Db 6701 ACAACATATCAAGAAAAAGTATCGTATCCAGAGGGGACCAAGGAGGCAATTTGTTTCAA 6760  
QY 2245 TAGAAAAAATAGAAAAATATAGACAAGCACATTTGTAATAGTAGAGCAAAATGGAATG 2304  
Db 6761 TAGAAAAAATAGAAAAATATAGACAAGCACATTTGTAATAGTAGAGCAAAATGGAATG 6820  
QY 2305 CCACTTTTAAACAGATATGTAACAAATTAAGAGAACATTTGGAATTAATTAACAAATTA 2364  
Db 6821 ACCTTTTAAACAGATATGTAACAAATTAAGAGAACATTTGGAATTAATTAACAAATTA 6880  
QY 2365 TCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTTGAAGCAAGTTTAAATTTTGGAG 2424  
Db 6881 TCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTTGAAGCAAGTTTAAATTTTGGAG 6940  
QY 2425 GGAATTTTCTACTGTAATTTCAACAACTGTTTAAATAGTACTGTTTAAATAGTACTT 2484  
Db 6941 GGAATTTTCTACTGTAATTTCAACAACTGTTTAAATAGTACTGTTTAAATAGTACTT 7000  
QY 2485 GGAATTTTCTACTGTAATTTCAACCTGGAAGAAAGTGAACAAATCACCCTCCATGAGAA 2544  
|||||

Db 7001 GGAATTTTCTACTGTAATTTCAACCTGGAAGAAAGTGAACAAATCACCCTCCATGAGAA 7060  
QY 2545 TAAACAAATTTTAAACATGTGCGACAGAAAGTGAAGAAAGCAATGTAATGCCCCCATCA 2604  
Db 7061 TAAACAAATTTTAAACATGTGCGACAGAAAGTGAAGAAAGCAATGTAATGCCCCCATCA 7120  
QY 2605 GTGGACAAATTTAGATGTTTCAATCAATATTAAGTGGCTGCTATTTAAACAGAGATGTGTA 2664  
Db 7121 GCGACAAATTTAGATGTTTCAATCAATATTAAGTGGCTGCTATTTAAACAGAGATGTGTA 7180  
QY 2665 ATAAACAATAGGCTCCGAGATCTTCAGACCTTGAGAGAGCCATATAGAGGCAATTTGA 2724  
Db 7181 ATAAACAATAGGCTCCGAGATCTTCAGACCTTGAGAGAGCCATATAGAGGCAATTTGA 7240  
QY 2725 GAAGTAAATTTATATATATTAAGTATGTAATTAATTAACCATATAGAGTATGACCCACCA 2784  
Db 7241 GAAGTAAATTTATATATATTAAGTATGTAATTAATTAACCATATAGAGTATGACCCACCA 7300  
QY 2785 AGCCAAAGAAAGAGTGTGTCAGAGAGAAAAAAGACAGTGGGAATAGAGCTTTGTTCC 2844  
Db 7301 AGCCAAAGAAAGAGTGTGTCAGAGAGAAAAAAGACAGTGGGAATAGAGCTTTGTTCC 7360  
QY 2845 TTGGGTTCTTTGGAGCAGACAGAGAAAGCACTATGGGCTGCACGTCAATGAGCTACGCTAC 2904  
Db 7361 TTGGGTTCTTTGGAGCAGACAGAGAAAGCACTATGGGCTGCACGTCAATGAGCTACGCTAC 7420  
QY 2905 AGGCCAGACATTTATGTTGATATATGATGTCAGCAGCAGAACTTTGCTGAGGGCTATTG 2964  
Db 7421 AGGCCAGACATTTATGTTGATATATGATGTCAGCAGCAGAACTTTGCTGAGGGCTATTG 7480  
QY 2965 AGGCCAGACATTTGTTGCAACATCAGACATCTGGGGCATCAAAAGCTCCAGCAGAGAA 3024  
Db 7481 AGGCCAGACATTTGTTGCAACATCAGACATCTGGGGCATCAAAAGCTCCAGCAGAGAA 7540  
QY 3025 TCTGCTGTGGGAAAGTATCTTAAGATATCAACAGCTCTGGGGATTTTGGGCTTGTCTG 3084  
Db 7541 TCTGCTGTGGGAAAGTATCTTAAGATATCAACAGCTCTGGGGATTTTGGGCTTGTCTG 7600  
QY 3085 GAAACATCATTTGTCACACATGCTGTGCTGTAATGCTATGTTGAGTAAATTAATCTCTG 3144  
Db 7601 GAAACATCATTTGTCACACATGCTGTGCTGTAATGCTATGTTGAGTAAATTAATCTCTG 7660  
QY 3145 AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAGAAATTTAACATTTACACAA 3204  
Db 7661 AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAGAAATTTAACATTTACACAA 7720  
QY 3205 GCTTAATACCTCTTAATTAATGAAGATCCGAAACCAAGCAAGAAATGAACAAGAAAT 3264  
Db 7721 GCTTAATACCTCTTAATTAATGAAGATCCGAAACCAAGCAAGAAATGAACAAGAAAT 7780  
QY 3265 TATTTGAATTTAGATTAATGAGCAAGTTTGTGAATTTGTTTAACATTAACAAATTTGCTGT 3324  
Db 7781 TATTTGAATTTAGATTAATGAGCAAGTTTGTGAATTTGTTTAACATTAACAAATTTGCTGT 7840  
QY 3325 GGTATATTAATTAATTTAT 3384  
Db 7841 GGTATATTAATTAATTTAT 7900  
QY 3385 CTGTACTTCTATATAGTAAT 3444  
Db 7901 CTGTACTTCTATATAGTAAT 7960  
QY 3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCGCCAGAAAGATATGAAGAAAGATGAGAGAG 3504  
Db 7961 ACCTCCCAATCCCGAGGGGACCCGACAGGCGCCAGAAAGATATGAAGAAAGATGAGAGAG 8020  
QY 3505 GAGACAGACAGATCCATTCAT 3564  
Db 8021 GAGACAGACAGATCCATTCAT 8080  
QY 3565 TGGGAGGCTGTGCTCTTTAGCTACCAAGGCTTTGAGAGACTTACTCTTATATATATATAT 3624  
Db 8081 TGGGAGGCTGTGCTCTTTAGCTACCAAGGCTTTGAGAGACTTACTCTTATATATATATATAT 8140  
|||||



```

QY 3625 GGATGTGGAAGTCTGTGGAGCAGGGGTGGGAGCCCTCAATATGTTGGATCTCC 3684
    |||||||
Db 8141 GGATGTGGAAGTCTGTGGAGCAGGGGTGGGAGCCCTCAATATGTTGGATCTCC 8200
    |||||||
QY 3685 TACAGTATTTGAGTGCAGCACTAAGAAATAGTGTCTTAACCTGCTCATATGCCACAGCA 3744
    |||||||
Db 8201 TACATATTTGGAGTGCAGCACTAAGAAATAGTGTCTTAACCTGCTCATATGCCACAGCA 8260
    |||||||
QY 3745 TAGCAGTATGCTGAGGGGACAGATAGGTTATAGAAATTAACAGCAGCTTATAGAGCTA 3804
    |||||||
Db 8261 TAGCAGTATGCTGAGGGGACAGATAGGTTATAGAAATTAACAGCAGCTTATAGAGCTA 8320
    |||||||
QY 3805 TTCCGCACATACCTTGAAGAAATTAAGACAGGGCTTGAAGAGTTTGTATATAGATGGGT 3864
    |||||||
Db 8321 TTCCGCACATACCTTGAAGAAATTAAGACAGGGCTTGAAGAGTTTGTATATAGATGGGT 8380
    |||||||
QY 3865 GGCAAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGGAAAGAAATGAGAGCA 3924
    |||||||
Db 8381 GGCAAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGGAAAGAAATGAGAGCA 8440
    |||||||
QY 3925 GCTGAGCCAGCAGCAGATGGGTGGGAGCAGTATCTGAGATCTAGA 3971
    |||||||
Db 8441 GCTGAGCCAGCAGCAGATGGGTGGGAGCAGTATCTGAGATCTAGA 8487
    |||||||

RESULT 14
AAN60128
ID AAN60128 standard; DNA; 3156 BP.
AC AAN60128;
XX
XX 25-MAR-2003 (updated)
DT 26-JUN-1991 (first entry)
XX
XX Sequence of the envelope gene of the HTLV-III proviral genome
DE (HXB-3).
XX
XX AIDS: HIV; LAV; HTLV-III: vaccine; antibody; epitope; antigen;
KM diagnosis; ss.
XX
XX HTLV-III.
OS
XX
XX Key Location/Qualifiers
FH CDS 478..3048
FT /*tag= a
FT
XX
XX EPI99301-A.
PN
XX
XX 29-OCT-1986.
PD
XX
XX 18-APR-1986; 86EP-0105371.
PF
XX
XX 19-APR-1985; 85US-0725021.
PR
XX
XX (HOPE ) HOFFMANN-LA ROCHE AG.
PA (USGO ) US GOVERNMENT.
PA (HEAL-) DEPT HEALTH & HUMAN SERV.
PA (USHS ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wongstaal FY;
PI MPI. 1986-286067/44.
DR P-PSDB; AAN60131.
XX
XX Envelope protein of acquired immune deficiency syndrome virus -
PT useful for improved testing of human blood for antibodies against
PT virus and as antigen for vaccines
XX
XX Claim 35; Fig 1; 46pp; English.
XX
XX An expression vector contg. a gene (AAN60128) coding for an envelope
CC protein of an AIDS virus, and the envelope protein of an AIDS virus

```

```

CC (AAN60131) are claimed. The vector is pref. a member of the pEV/env
CC family, e.g. pEV1.2 or 3/env 44-640 or 205-640.
CC (updated on 25-MAR-2003 to correct PA field.)
CC (updated on 25-MAR-2003 to correct PI field.)
XX
XX
SQ Sequence 3156 BP; 1095 A; 536 C; 765 G; 760 T; 0 other;
Query Match 41.7%; Score 2598.2; DB 7; Length 3156;
Best Local Similarity 98.4%; Pred. No. 1,9e-304;
Matches 2637; Conservative 0; Mismatches 38; Indels 6; Gaps 1;
QY 1291 CCATAGAGTGAAGAGAGAGTATCAGCCTTGTGAGATGGGGTGAAGAGGACCA 1350
    |||||||
Db 476 CAATAGAGTGAAGAGAGAGTATCAGCCTTGTGAGATGGGGTGAAGAGGACCA 535
    |||||||
QY 1351 TGCCTCTTGGATTTATGATCTGTAGTCTACAGAAATTTGGGTCAACCTCTATT 1410
    |||||||
Db 536 TGCCTCTTGGATTTATGATCTGTAGTCTACAGAAATTTGGGTCAACCTCTATT 595
    |||||||
QY 1411 ATGGGCTACCTGTGTGAAGAGAGCAACACACCTCTATTGTGCATCAGATGCTAAG 1470
    |||||||
Db 596 ATGGGCTACCTGTGTGAAGAGAGCAACACACCTCTATTGTGCATCAGATGCTAAG 655
    |||||||
QY 1471 CATATGATACAGAGGTACATATCTTTGGGCCACACATGCTGTACCCACAGACCCA 1530
    |||||||
Db 656 CATATGATACAGAGGTACATATCTTTGGGCCACACATGCTGTACCCACAGACCCA 715
    |||||||
QY 1531 ACCCAGAGAGTACTATTGTTAAATGTGACAGAAATTTTACATGTGGAAATGACA 1590
    |||||||
Db 716 ACCCAGAGAGTACTATTGTTAAATGTGACAGAAATTTTACATGTGGAAATGACA 775
    |||||||
QY 1591 TGGTGAACAGATGATAGATATTAATCAGTTTGGATCAAAAGCCATAAGCATGAG 1650
    |||||||
Db 776 TGGTGAACAGATGATAGATATTAATCAGTTTGGATCAAAAGCCATAAGCATGAG 835
    |||||||
QY 1651 TAAATTTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATATCTAATA 1710
    |||||||
Db 836 TAAATTTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATATCTAATA 895
    |||||||
QY 1711 CCAATAGTATGTAAGCGGAGCAATGATTAATGGAGAAAGAGAGATTAATAAACTGCTTTCA 1770
    |||||||
Db 896 CCAATAGTATGTAAGCGGAGCAATGATTAATGGAGAAAGAGAGATTAATAAACTGCTTTCA 955
    |||||||
QY 1771 ATATCAGCAGCAAGTAAAGAGATTAAGTGCAGAAAGAAATATGCTTTTAAACTTG 1830
    |||||||
Db 956 ATATCAGCAGCAAGTAAAGAGATTAAGTGCAGAAAGAAATATGCTTTTAAACTTG 1015
    |||||||
QY 1831 ATATAGTACCAATAGATTA-----TACCAGCTATAGTGTGATTAAGTGTGACACCTGAG 1884
    |||||||
Db 1016 ATATATATACCAATAGATTAAGTGTGACACCTATAGCTGTGACAGTGTGATTAAGTGTGACACCTGAG 1075
    |||||||
QY 1885 TCATTACAGAGCGCTGTCCAAAGGTATGCTTTGACACCAATTCCTTATCTATTTGTGCCC 1944
    |||||||
Db 1076 TCATTACAGAGCGCTGTCCAAAGGTATGCTTTGACACCAATTCCTTATCTATTTGTGCCC 1135
    |||||||
QY 1945 CGGCTGTTTGGCTTCAAAATGTAATTAATTAAGCTTCAATGAAGAGAGCATGTA 2004
    |||||||
Db 1136 CGGCTGTTTGGCTTCAAAATGTAATTAATTAAGCTTCAATGAAGAGAGCATGTA 1195
    |||||||
QY 2005 CAAATGTAGCAGCAGTACATATGTACACATGAGCAATGAGCAGTATATCACTCAACTGCG 2064
    |||||||
Db 1196 CAAATGTAGCAGCAGTACATATGTACACATGAGCAATGAGCAGTATATCACTCAACTGCG 1255
    |||||||
QY 2065 TGTTAATGTGAGCTGTACAGAGAGATGTATTAATAGATCTGCCAATTTTCACAGACA 2124
    |||||||
Db 1256 TGTTAATGTGAGCTGTACAGAGAGATGTATTAATAGATCTGCCAATTTTCACAGACA 1315
    |||||||
QY 2125 ATGCTAAACCAATATAGTACAGCTGAGACACATCTGTGAATTTATGTATGAACAGCCA 2184
    |||||||
Db 1316 ATGCTAAACCAATATAGTACAGCTGAGACACATCTGTGAATTTATGTATGAACAGCCA 1375
    |||||||
QY 2185 ACAACAATACAGAAAGATATCCGTATTCAGAGGGAGCCAGGAGACATTTGTACAA 2244
    |||||||

```

```

Db 1376 ACAACATACAGAAAAAAATCCGTATCCAGAGGGGACGAGGAGACATTTGTTACAA 1435
Oy 2245 TAGGAAAAATAGAAAATATGAGACAGACATTTGTAACATTAGTAGACAAATGGAATG 2304
Db 1436 TAGGAAAAATAGAAAATATGAGACAGACATTTGTAACATTAGTAGACAAATGGAATG 1495
Oy 2305 CCACTTTAAACAGATAGCTAGCAAAATTTAGAGAACAAATTTGAAATTAATTAACAAATTA 2364
Db 1496 CCACCTTTAAACAGATAGCTAGCAAAATTTAGAGAACAAATTTGAAATTAATTAACAAATTA 1555
Oy 2365 TCTTTAAGCAATCTCTGAGAGAGGAGCCAGAAATTTGTAGCCACAGTTTAAATTTGAGAG 2424
Db 1556 TCTTTAAGCAATCTCTGAGAGAGGAGCCAGAAATTTGTAGCCACAGTTTAAATTTGAGAG 1615
Oy 2425 GGGATTTTCTACTGTATTCACACAACTGTTTAATAGTACTGTTTAAATGTTACTT 2484
Db 1616 GGGATTTTCTACTGTATTCACACAACTGTTTAATAGTACTGTTTAAATGTTACTT 1675
Oy 2485 GGAAGTACTGAAGGGTCAAAATTAACACTGAAGAGAGTGACACATCCACTCCCATGACAGAA 2544
Db 1676 GGAAGTACTGAAGGGTCAAAATTAACACTGAAGAGAGTGACACATCCACTCCCATGACAGAA 1735
Oy 2545 TAAACACATTTATTAACATGTCGAGAGAGAGTAAGAAAAAGCAATGTATGCCCCCTCCATCA 2604
Db 1736 TAAACACATTTATTAACATGTCGAGAGAGAGTAAGAAAAAGCAATGTATGCCCCCTCCATCA 1795
Oy 2605 GTGACACAAATAGATGTTCAATCAATATTTACTGGGCTGCTATTAACAGAGATGGTGTA 2664
Db 1796 GCGACACAAATAGATGTTCAATCAATATTTACTAGGGCTGCTATTAACAGAGATGGTGTA 1855
Oy 2665 ATAAACACATGTTGGTCCGAGATCTTCAGAGAGGCGGATGTAGGAGCAATTTGGA 2724
Db 1856 ATAAACACAAAGGGTCCGAGATCTTCAGAGAGGCGGATGTAGGAGCAATTTGGA 1915
Oy 2725 GAAGTGAATTTATTAATATTAAGTAGTAATTAAGCAATTTAGAGTATGACCCACCA 2784
Db 1916 GAAGTGAATTTATTAATATTAAGTAGTAATTAAGCAATTTAGAGTATGACCCACCA 1975
Oy 2785 AGGCAAGAGAGAGAGTGGTGACAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCC 2844
Db 1976 AGGCAAGAGAGAGTGGTGACAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCC 2035
Oy 2845 TTGGGTTCTTGGGAGCAGCAGAGAAAGCACTATGGCTGCACGTCAATGACGCTGACGCTAC 2904
Db 2036 TTGGGTTCTTGGGAGCAGCAGAGAAAGCACTATGGGCGCAGCCTCAATGACGCTGACGCTAC 2095
Oy 2905 AGGCGACAGCAATTTGTTCTGATATAGTGTGAGCAGCAGCAAAATTTGCTGAGGGCTATTG 2964
Db 2096 AGGCGACAGCAATTTGTTCTGATATAGTGTGAGCAGCAGCAAAATTTGCTGAGGGCTATTG 2155
Oy 2965 AGGCGCAACAGACATCTGTTGCAACTCAGAGTCTGGGGCATTCAACAGCTCCAGGCAAGAA 3024
Db 2156 AGGCGCAACAGACATCTGTTGCAACTCAGAGTCTGGGGCATTCAACAGCTCCAGGCAAGAA 2215
Oy 3025 TCCCTGGCTGTGGAAGATACCTTAAGATCAACAGCTCTCTGGGATTTGGGGTGTCTGTG 3084
Db 2216 TCCCTGGCTGTGGAAGATACCTTAAGATCAACAGCTCTCTGGGATTTGGGGTGTCTGTG 2275
Oy 3085 GAAACTCATTTGACACACTCTGCTGCTTGAATGCTAGTGTGAGTAATTAATCTGTG 3144
Db 2276 GAAACTCATTTGACACACTCTGCTGCTTGAATGCTAGTGTGAGTAATTAATCTGTG 2335
Oy 3145 AACAGATTTGGAATTAAGATGACGTGATGAGTGGGACAGAGAAATTTAACAAATTAACAA 3204
Db 2336 AACAGATTTGGAATTAAGATGACGTGATGAGTGGGACAGAGAAATTTAACAAATTAACAA 2395
Oy 3205 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATAAACAGAAAT 3264
Db 2396 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATAAACAGAAAT 2455
Oy 3265 TATTTGAATTTAGTAATTAAGGCAAGTTTGTGAATTTGTTTAACATAACAAATTTGGGCT 3324
Db 2456 TATTTGAATTTAGTAATTAAGGCAAGTTTGTGAATTTGTTTAACATAACAAATTTGGGCT 2515

```

```

Oy 3325 GGTATTAATAAATTAATTCATAATAGTAGTAGAGGCTTGGTAGTTAAGAAATAGTATTG 3384
Db 2516 GGTATTAATAAATTAATTCATAATAGTAGTAGAGGCTTGGTAGTTAAGAAATAGTATTG 2575
Oy 3385 CTGTACTTTCTATAGTGAATAGAGTGTAGGAGGATATTCACATTTATGCTTTAGACCC 3444
Db 2576 CTGTACTTTCTGTAGTGAATAGAGTGTAGGAGGATATTCACATTTATGCTTTAGACCC 2635
Oy 3445 ACCGCCAATCCCGAGGGGACCCGACAGGCCCGGAAGGAATTAAGAAAGAGGTGACAGA 3504
Db 2636 ACCGCCAATCCCGAGGGGACCCGACAGGCCCGGAAGGAATTAAGAAAGAGGTGACAGA 2695
Oy 3505 GAGACAGAGACAGATCCATTCATTTAGTAAGACGATCCCTTATGACTTATCGGACAGATC 3564
Db 2696 GAGACAGAGACAGATCCATTCATTTAGTAAGACGATCCCTTATGACTTATCGGACAGATC 2755
Oy 3565 TCGGAGGCTGTGCTCTTTCAGCTACACACCGCTTGAGAGACTTACTTGAATTTAACA 3624
Db 2756 TCGGAGGCTGTGCTCTTTCAGCTACACACCGCTTGAGAGACTTACTTGAATTTAACA 2815
Oy 3625 GGATTTGGAATCTTGGAGAGCAGGGGTGGGAAGCCCTCAATATTTGGGAATCTCC 3684
Db 2816 GGATTTGGAATCTTGGAGAGCAGGGGTGGGAAGCCCTCAATATTTGGGAATCTCC 2875
Oy 3685 TACAGTATTTGGAGTCAGAGAACATAAAGATAGTGTGTTAACTGCTCAATGCCACAGCA 3744
Db 2876 TACAGTATTTGGAGTCAGAGAACATAAAGATAGTGTGTTAACTGCTCAATGCCACAGCA 2935
Oy 3745 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTTCAAGCAGCTTATAGAGCTA 3804
Db 2936 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAATTTCAAGAACTTATAGAGCTA 2995
Oy 3805 TTGCCACATCTCTAGAAAGATAAGACAGGCGCTTGAAGAGATTTGCTATTAAGATGGGT 3864
Db 2996 TTGCCACATCTCTAGAAAGATAAGACAGGCGCTTGAAGAGATTTGCTATTAAGATGGGT 3055
Oy 3865 GGCAAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAAGGAAAAAGATGACAGA 3924
Db 3056 GGCAAGTGTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAAGGAAAAAGATGACAGA 3115
Oy 3925 GCTGAGCCAGCAGCAGATGGGCTGGAGCAGTATCTGAGA 3965
Db 3116 GCTGAGCCAGCAGCAGATGGGCTGGAGCAGCAGTCTGAGA 3156

RESULT 15
AAH76385
ID AAH76385 standard; DNA; 8932 BP.
XX
AC AAH76385;
XX
AC 30-NOV-2001 (first entry)
XX
DE Nucleotide sequence of HIV-1 isolate BH10.
XX
KW HIV-1; gp120; BH10; vaccine; immunization; ds.
XX
OS Human immunodeficiency virus type 1.
XX
PN US6268484-B1.
XX
PD 31-JUL-2001.
XX
PE 30-JUL-1998; 98US-0124900.
XX
PR 07-JUN-1995; 95US-0478536.
PR 19-APR-1995; 95WO-EP01481.
PA (POLY-) POLYMUN SCI IMMUNOBIOLOGISCHE FORSCHUNG.
XX
PI Kätlinger H, Buchacher A, Ernst W, Ballaun C, Purtscher M;
PI Trkola A, Predl R, Schmätzl C, Klima A, Steindl F, Muster T;

```



QY 3085 GAAACCTATTTTCACCACTGCTGTGCTTGAATGCTAGTGAATTAATTCCTGG 3144  
|||||  
Db 7378 GAAACTCATTTTCACCACTGCTGTGCTTGAATGCTAGTGAATTAATTCCTGG 7437  
|||||  
QY 3145 AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAAATTAACATTTACCAA 3204  
|||||  
Db 7438 AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAAATTAACATTTACCAA 7497  
|||||  
QY 3205 GCTTAATACACCTCTTAATTTGAAGATTCGCAAAACAGCAAGAAATGAACAGAAAT 3264  
|||||  
Db 7498 GCTTAATACACCTCTTAATTTGAAGATTCGCAAAACAGCAAGAAATGAACAGAAAT 7557  
|||||  
QY 3265 TATTGGAATTTAGTAAATGGGCAAGTTTGTGAATTTGTTAAATTAACAAATTTGGCTGT 3324  
|||||  
Db 7558 TATTGGAATTTAGTAAATGGGCAAGTTTGTGAATTTGTTAAATTAACAAATTTGGCTGT 7617  
|||||  
QY 3325 GGTATATAAATTAATTCATTAATGATAGTAGAGGCTTGTAGGTTTAAGAAATAGTTTGTG 3384  
|||||  
Db 7618 GGTATATAAATTAATTCATTAATGATAGTAGAGGCTTGTAGGTTTAAGAAATAGTTTGTG 7677  
|||||  
QY 3385 CTGTACTTCTATAGTAATAGATAGGATATTCACCATTTTCCTTCAGACCC 3444  
|||||  
Db 7678 CTGTACTTCTATAGTAATAGATAGGATATTCACCATTTTCCTTCAGACCC 7737  
|||||  
QY 3445 ACCTCCCATCCCGAGGGGACCCGACAGGCCGGAAGGAATAGAAGAGAGTGAGAGAGA 3504  
|||||  
Db 7738 ACCTCCCATCCCGAGGGGACCCGACAGGCCGGAAGGAATAGAAGAGAGTGAGAGAGA 7797  
|||||  
QY 3505 GAGACAGAGACAGATCCATTCGATTAAGTGAACGATCCTTACGACTTATCTGGAGCATC 3564  
|||||  
Db 7798 GAGACAGAGACAGATCCATTCGATTAAGTGAACGATCCTTACGACTTATCTGGAGCATC 7857  
|||||  
QY 3565 TGGGAGGCCGTGGCTCTTACGATACCAAGCTTGAAGACTTACTCTGATTTGTAACGA 3624  
|||||  
Db 7858 TGGGAGGCCGTGGCTCTTACGATACCAAGCTTGAAGACTTACTCTGATTTGTAACGA 7917  
|||||  
QY 3625 GGATTTGTGAACCTTGTGGAGCAGGAGGGTGGGAAAGCCCTCAAAATATTTGGTGAATCTCC 3684  
|||||  
Db 7918 GGATTTGTGAACCTTGTGGAGCAGGAGGGTGGGAAAGCCCTCAAAATATTTGGTGAATCTCC 7977  
|||||  
QY 3685 TACAGTATTTGAGCTCAGGAACCTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCA 3744  
|||||  
Db 7978 TACAGTATTTGAGCTCAGGAACCTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCTA 8037  
|||||  
QY 3745 TACAGTATTTGAGGAGGAGCAGATAGGTTATAGAGTATTAACAAGCAGCTTATAGAGCTA 3804  
|||||  
Db 8038 TACAGTATTTGAGGAGGAGCAGATAGGTTATAGAGTATTAACAAGCAGCTTATAGAGCTA 8097  
|||||  
QY 3805 TTGGCCACATACCTAGAGAATTAAGACAGGGCTTGAAGGATTTTCTATAAGATGGT 3864  
|||||  
Db 8098 TTGGCCACATACCTAGAGAATTAAGACAGGGCTTGAAGGATTTTCTATAAGATGGT 8157  
|||||  
QY 3865 GGCAGTGTGCACAAAAGTACTGTGATTTGATGGCTGCTTAAGGAAAGATGAGAGCA 3924  
|||||  
Db 8158 GGCAGTGTGCACAAAAGTACTGTGATTTGATGGCTGCTTAAGGAAAGATGAGAGCA 8217  
|||||  
QY 3925 GCTGAGCCAGCAGAGATGGGTTGGGAGCAGTATCTCGAGATCTAGA 3971  
|||||  
Db 8218 GCTGAGCCAGCAGAGATGGGTTGGGAGCAGATCTCGAGATCTAGA 8264  
|||||

Search completed: September 17, 2003, 08:05:26  
Job time : 1488 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:40:41 : Search time 320 Seconds  
(without alignments)  
8591.799 Million cell updates/sec

Title: US-09-913-159A-10

Perfect score: 6229

Sequence: 1 ctgacgcgcctctgacgcgc.....attcccccgaagtcgcac 6229

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents, NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/6D.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2676.2	43.0	9709	2	US-08-188-583-5
2	2676.2	43.0	9709	3	US-08-388-353-1
3	2676.2	43.0	9709	4	US-08-488-551B-1
4	2676.2	43.0	9709	4	US-09-309-572-15
5	2636.6	42.3	15581	3	US-08-646-538-35
6	2636.6	42.3	15581	3	US-09-503-222-35
7	2597.8	41.7	8933	3	US-09-124-900-1
8	2597.8	41.7	8933	3	US-08-463-210-4
9	2597.8	41.7	8933	3	US-09-620-958A-3
10	2597.8	41.7	8933	3	US-09-620-958A-4
11	2597.8	41.7	8933	3	US-09-620-958A-9
12	2593	41.6	9719	4	US-09-700-304-1
13	2590.8	41.6	2945	6	5462872-1
14	2558.4	41.1	12479	4	US-09-318-138-13
15	2558.4	41.1	12494	3	US-08-935-312-13
16	2558.4	41.1	12494	3	US-08-848-760B-33
17	2554.6	41.0	3563	3	US-08-463-210-6
18	2486.6	39.9	2571	2	US-07-916-098A-1
19	2411.6	38.7	2644	1	Sequence 6, Appli
20	2303.4	37.0	2531	3	US-08-472-240A-9
21	2205.6	35.4	3807	3	US-07-956-483-18
22	2205.6	35.4	3807	1	US-08-022-835-5
23	2205.6	35.4	3807	2	US-08-388-809-5
24	2199.4	35.3	2730	4	US-08-647-714-5
25	2195.4	35.2	2696	4	US-08-728-122-1
26	2141.6	34.4	9737	2	US-09-325-131B-1
27	2141.6	34.4	9737	4	US-08-944-449-7
					Sequence 7, Appli

28	2118.8	34.0	2552	2	US-08-448-603A-27	Sequence 27, Appl
29	2118.8	34.0	2552	3	US-09-134-075-27	Sequence 27, Appl
30	2118.8	34.0	2552	4	US-09-492-739-27	Sequence 27, Appl
31	2113	33.9	9746	1	US-08-022-835-3	Sequence 3, Appli
32	2113	33.9	9746	2	US-08-388-809-3	Sequence 3, Appli
33	2113	33.9	9746	1	US-08-647-714-3	Sequence 3, Appli
34	2111.8	33.9	4527	2	US-08-944-449-8	Sequence 8, Appli
35	2111.8	33.9	4527	4	US-09-353-362-8	Sequence 8, Appli
36	2101.8	33.7	9739	1	US-08-022-835-1	Sequence 1, Appli
37	2101.8	33.7	9739	1	US-08-388-809-1	Sequence 1, Appli
38	2101.8	33.7	9739	2	US-08-647-714-1	Sequence 1, Appli
39	2091.2	33.6	6474	3	US-08-651-472-66	Sequence 66, Appl
40	2091.2	33.6	6474	3	US-08-358-928-66	Sequence 66, Appl
41	2090.6	33.6	6926	3	US-08-651-472-69	Sequence 69, Appl
42	2090.6	33.6	6926	3	US-08-358-928-69	Sequence 69, Appl
43	2088.2	33.5	6573	2	US-08-448-603A-29	Sequence 29, Appl
44	2088.2	33.5	2573	3	US-09-134-075-29	Sequence 29, Appl
45	2088.2	33.5	2573	4	US-09-492-739-29	Sequence 29, Appl

## ALIGNMENTS

RESULT 1  
US-08-188-583-5  
Sequence 5, Application US/08188583  
Patent No. 5851813  
GENERAL INFORMATION:  
APPLICANT: Destroiers, Ronald C.  
TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: WordPerfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188, 583  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/27,494  
FILING DATE: July 9, 1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/551,945  
FILING DATE: July 12, 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: Reg. No. 5851813 29,066  
REFERENCE/DOCKET NUMBER: 00246/079002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9709  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-188-583-5  
Query Match 43.0%, Score 2676.2, DB 2, Length 9709;  
Best Local Similarity 99.9%, Pred. No. 0;  
Matches 2678; Conservative 0; Mismatches 3; Indels 0;





```

; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-388-353-1

Query Match          43.0%      Score 2676.2;   DB 3;   Length 9709;
Best Local Similarity 99.9%      Pred. No. 0;
Matches 2678;   Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

```

QY	1231	CCATGACAGCTGAAGAGAAAGTATCATGACACTGTGTGAGATGGGGGTGGAAATGGGGCACCA	1350
Db	6219	CAATGAGAGTAAAGGAAAGTATGACGACTTGTGGAGATGGGGGTGGAAATGGGGCACCA	6278
QY	1351	TGCTCCCTGGGATATTTGATTCGTCTAGTGCCTACGAAAAATTTGGGGTCCACGGCTATT	1410
Db	6279	TGCTCCTTGGGATATTTGATTCGTCTAGTGCCTACGAAAAATTTGGGGTCCACAGCTATT	6338
QY	1411	ATGGGGTACCTGTGTGGAAGGAAAGCAACACCACTTATTTTGTGCATCAGATGCTAAAG	1470
Db	6339	ATGGGGTACCTGTGTGGAAGGAAAGCAACACCACTCTTATTTTGTGCATCAGATGCTAAAG	6398
QY	1471	CATATGATACAGAGGTACATATATTTTGGGGCACACAGCCCTGTGTACCCACACACCCCA	1530
Db	6399	CATATGATACAGAGGTACATATATTTTGGGGCACACAGCCCTGTGTACCCACAGCCCCA	6458
QY	1531	ACCCACAAGATGTATTTGGTAAATGTGACAGAAATTTTACATGTGGAATAATGACA	1590
Db	6459	ACCCACAAGATGTATTTGGTAAATGTGACAGAAATTTTACATGTGGAATAATGACA	6518
QY	1591	TGCTGGAACAGATGCATGAGATATATCAGTTTATGGGATCAAAAGCTTAAAGCATGTG	1656
Db	6519	TGCTGGAACAGATGCATGAGATATATCAGTTTATGGGATCAAAAGCTTAAAGCATGTG	6578
QY	1651	TAAATTTAACCCACTGCTGTGTGTTTAAAGTGCACATGTTTGAAGAAATATCTATA	1710
Db	6579	TAAATTTAACCCACTGCTGTGTGTTTAAAGTGCACATGTTTGAAGAAATATCTATA	6638
QY	1711	CCAATAGTAGTAGGGGGAATGATTAATGAGAAAAGAGATATAAAATCTCTTTTCA	1770
Db	6639	CCAATAGTAGTAGGGGGAATGATTAATGAGAAAAGAGATATAAAATCTCTTTTCA	6698
QY	1771	ATATCAGCACAGCATTAAGAGATTAAGTGTGCAAAAGATATGCATCTTTTATTAACCTTG	1830
Db	6699	ATATCAGCACAGCATTAAGAGATTAAGTGTGCAAAAGATATGCATCTTTTATTAACCTTG	6758
QY	1831	ATTATAGTACCATTTGATATATACAGCTTATAGTTTGAATGATTTGTAACCCCTCAGCATTA	1890
Db	6759	ATTATAGTACCATTTGATATATACAGCTTATAGTTTGAATGATTTGTAACCCCTCAGCATTA	6818
QY	1891	CACAGGCCGTGCCAAAGGTATCCTTTGAGCCAAATCCCATCATTTATTTGTGCCCGGCTG	1950
Db	6819	CACAGGCCGTGCCAAAGGTATCCTTTGAGCCAAATCCCATCATTTATTTGTGCCCGGCTG	6878
QY	1951	GTTTTGGATTTCTAAATGTATATATAAGAGTTTCAATGAAGAAAGGACCAATGACAAATG	2010
Db	6879	GTTTTGGATTTCTAAATGTATATATAAGAGTTTCAATGAAGAAAGGACCAATGACAAATG	6938
QY	2011	TCAGACACAGTCAATGTATACATATGAATCAGGCCAGTAGTCAATCAATCACTGCGTTAA	2070
Db	6939	TCAGACACAGTCAATGTATACATATGAATCAGGCCAGTAGTCAATCAATCACTGCGTTAA	6998
QY	2071	ATGGCAGTCTAGCAGAAAGATGTAGTAAATTTAGATCTGCCAATTTTCCACAGACAATGCTA	2130
Db	6999	ATGGCAGTCTAGCAGAAAGATGTAGTAAATTTAGATCTGCCAATTTTCCACAGACAATGCTA	7058
QY	2131	AAACCATATATGTACAGCTGAAACACATCTGTAGAATTTAATTTGTACAAAGACCCAACA	2190
Db	7059	AAACCATATATGTACAGCTGAAACACATCTGTAGAATTTAATTTGTACAAAGACCCAACA	7118
QY	2191	ATTACAGAAAAGATATCCGTATCCAGAGGGGACACAGGAGACATTTTGTTCATATAGGAA	2250



: FILING DATE: 14-FEB-1994
: APPLICATION NUMBER: PM4002 (AU)
: FILING DATE: 21-FEB-1994
: APPLICATION NUMBER: PNO284 (AU)
: FILING DATE: 23-DEC-1994
: APPLICATION NUMBER: US 08/388,353
: FILING DATE: 14-FEB-1995
: APPLICATION NUMBER: PN3021/95
: FILING DATE: 17-MAY-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: FRANK S. DIGIGLIO
: REFERENCE/DOCKET NUMBER: 9606Z
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 742-4343
: TELEFAX: (516) 742-4366
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9709 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-488-551B-1

Query Match 43.0%; Score 2676.2; DB 3; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1291 CCATGAGAGTGAAGAGAGATATCAGCATTGTGGAGATGGGGGTGNAATGGGGCACCA 1350
Db 6219 CAATGAGAGTGAAGAGAGATATCAGCATTGTGGAGATGGGGGTGNAATGGGGCACCA 6278
QY 1351 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAATTGGGGTCAACCGTCTATT 1410
Db 6279 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAATTGGGGTCAACCGTCTATT 6338
QY 1411 ATGGGGTACCTGTGTGGAAGAGACACACACCTCTATTGTGTCATCAGATGCTAAG 1470
Db 6339 ATGGGGTACCTGTGTGGAAGAGACACACACCTCTATTGTGTCATCAGATGCTAAG 6398
QY 1471 CATATGATACAGAGTACATATATGTTTGGGCGACACATGCTGTACCCACAGACCCA 1530
Db 6399 CATATGATACAGAGTACATATATGTTTGGGCGACACATGCTGTACCCACAGACCCA 6458
QY 1531 ACCCAAGAAGTATGTTGTTAAATGTCAGAGAAAATTTTAACTGTGGGAAAATGACA 1590
Db 6459 ACCCAAGAAGTATGTTGTTAAATGTCAGAGAAAATTTTAACTGTGGGAAAATGACA 6518
QY 1591 TGGTAGAAGCAGATGAGGATATATCAGTTTATGGGATCAAAAGCCTTAAGCCATGTG 1650
Db 6519 TGGTAGAAGCAGATGAGGATATATCAGTTTATGGGATCAAAAGCCTTAAGCCATGTG 6578
QY 1651 TAAATTTAACCCCACTGTGTAGTTTAAAGTGCACATGATTTGNAAGATATATAATA 1710
Db 6579 TAAATTTAACCCCACTGTGTAGTTTAAAGTGCACATGATTTGNAAGATATATAATA 6638
QY 1711 CCAATAGTAGTAGGGGGGAATGATATGGAAGAAAGAGATATAAAAACTGCTTTTCA 1770
Db 6639 CCAATAGTAGTAGGGGGGAATGATATGGAAGAAAGAGATATAAAAACTGCTTTTCA 6698
QY 1771 ATATCAGCACAAGCATAGAGATTAAGTGTCAGAAAAGATATGCAATCTTTTATTAACCTG 1830
Db 6699 ATATCAGCACAAGCATAGAGATTAAGTGTCAGAAAAGATATGCAATCTTTTATTAACCTG 6758
QY 1831 ATATAGTACCAATAGATATAACAGCTATAGTTGATAGTTTAAACCTCAGTCATTA 1890
Db 6759 ATATAGTACCAATAGATATAACAGCTATAGTTGATAGTTTAAACCTCAGTCATTA 6818
QY 1891 CACAGGCTGTCCAAAGGTATCTTTGAGCCAATTTCCATACATTTATTGTGCCCCGGGTG 1950
Db 6819 CACAGGCTGTCCAAAGGTATCTTTGAGCCAATTTCCATACATTTATTGTGCCCCGGGTG 6878
QY 1951 GTTTTGCATTTCTAAATGTATATATAGAGCTTCAATGGAACGAGACATGTACAATG 2010

Db 6879 GTTTTGCATTTCTAAATGTATATATAGACCTTCAATGACAGACCATGTACAAATG 6938
QY 2011 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGATATCAACTCACTGCTGTAA 2070
Db 6939 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGATATCAACTCACTGCTGTAA 6998
QY 2071 ATGGCAGTCTAGCAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACAAATGCTA 2130
Db 6999 ATGGCAGTCTAGCAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACAAATGCTA 7058
QY 2131 AAACCATATATAGTACAGTGAACACATCTGTAGAAATTAATTTGACAGACCAACAACA 2190
Db 7059 AAACCATATATAGTACAGTGAACACATCTGTAGAAATTAATTTGACAGACCAACAACA 7118
QY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGAGAGCATTTGTTTCAATAGGAA 2250
Db 7119 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGAGAGCATTTGTTTCAATAGGAA 7178
QY 2251 AAATAGGAATATGAGACAGACACATTTGTAACATTAGTAGAGCAAAATGCAATGCCATT 2310
Db 7179 AAATAGGAATATGAGACAGACACATTTGTAACATTAGTAGAGCAAAATGCAATGCCATT 7238
QY 2311 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGGAATATATAAACAATATCTTTA 2370
Db 7239 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGGAATATATAAACAATATCTTTA 7298
QY 2371 AGCAATCCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAATTTGTGAGGGGAAT 2430
Db 7299 AGCAATCCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAATTTGTGAGGGGAAT 7358
QY 2431 TTTTCTACTGTATTAATTAACACAACTGTTTATAGTACTGTTTAAATAGTACTGGAATA 2490
Db 7359 TTTTCTACTGTATTAATTAACACAACTGTTTATAGTACTGTTTAAATAGTACTGGAATA 7418
QY 2491 CTGAAGGGTCAAAATTAACACTGTAAGAGAAATGACACATCACTCCCATGCAATTAAC 2550
Db 7419 CTGAAGGGTCAAAATTAACACTGTAAGAGAAATGACACATCACTCCCATGCAATTAAC 7478
QY 2551 AATTTATTAACATGTGGCAGGAGTAGGAAAAAGCAATGATCCCTCCATCAGTGGAC 2610
Db 7479 AATTTATTAACATGTGGCAGGAGTAGGAAAAAGCAATGATCCCTCCATCAGTGGAC 7538
QY 2611 AAATTAAGTGTTCATCAATATTTACTGGGCTGCTATTAAACAAGATGTTGGTATATACA 2670
Db 7539 AAATTAAGTGTTCATCAATATTTACTGGGCTGCTATTAAACAAGATGTTGGTATATACA 7598
QY 2671 ACAATGGGTCCGAGATCTTACAGACCTGAGAGAGCGATATGAGGACAATTTGGAAGATG 2730
Db 7599 ACAATGGGTCCGAGATCTTACAGACCTGAGAGAGCGATATGAGGACAATTTGGAAGATG 7658
QY 2731 AATTAATTAATTAATTAAGTATTAATAAATTTGAACCATTTAGAGTAGACCCACAAGGCAA 2790
Db 7659 AATTAATTAATTAATTAAGTATTAATAAATTTGAACCATTTAGAGTAGACCCACAAGGCAA 7718
QY 2791 AGAGAAAGTGGTGCAGAGAGAAAAAARAGCAATGGGAGATAGACCTTTGTCTTGGGT 2850
Db 7719 AGAGAAAGTGGTGCAGAGAGAAAAAARAGCAATGGGAGATAGACCTTTGTCTTGGGT 7778
QY 2851 TCTTGGGAGCAGCAGAGAGCACTATGGGCTCAGCTCAATAGCCTGAGGCTACAGGCCA 2910
Db 7779 TCTTGGGAGCAGCAGAGAGCACTATGGGCTCAGCTCAATAGCCTGAGGCTACAGGCCA 7838
QY 2911 GACAAATTTTGTCTGATATAGTGCAGCAGCAAAATTTCTGAAGGGCTATTGAGGGCG 2970
Db 7839 GACAAATTTTGTCTGATATAGTGCAGCAGCAAAATTTCTGAAGGGCTATTGAGGGCG 7898
QY 2971 AACAGCATCTGTGCAATCTCAGAGTCTGGGCGATCAACACACTCCAGGCAAGATCTGG 3030
Db 7899 AACAGCATCTGTGCAATCTCAGAGTCTGGGCGATCAACACACTCCAGGCAAGATCTGG 7958
QY 3031 CTGTGGAAGATPACTTAAGGATCAACAGCTCTGGGGATTTGGGGTTGCTGTGAAAC 3090

Db 7959 CTGTGGAAGATACCTAAAGATCAACAGCTCTGCGGGATTGGGGTTGCTCTGGAAGAAC 8018  
 QY 3091 TCATTTGCACACACTGCTGTGCTTGAATGCTAGTGGAGTAATTAATCTCTGGAACAGA 3150  
 Db 8019 TCATTTGCACACACTGCTGTGCTTGAATGCTAGTGGAGTAATTAATCTCTGGAACAGA 8078  
 QY 3151 TTTGGAATTAACATGACCTGTGAGTGGAGAGAGAAATTAACATTTACCAAGCTTTAA 3210  
 Db 8079 TTTGGAATTAACATGACCTGTGAGTGGAGAGAGAAATTAACATTTACCAAGCTTTAA 8138  
 QY 3211 TACACCTCTTAATGAGAAATCGCAAAACCGCAAGAAAGAAATGAACAAATATTATGG 3270  
 Db 8139 TACACCTCTTAATGAGAAATCGCAAAACCGCAAGAAAGAAATGAACAAATATTATGG 8198  
 QY 3271 AATTAGATTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 3330  
 Db 8199 AATTAGATTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 8258  
 QY 3331 TAAATTAATTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 3390  
 Db 8259 TAAATTAATTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 8318  
 QY 3391 TTTCTATAGTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 3450  
 Db 8319 TTTCTATAGTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 8378  
 QY 3451 CAATCCCGAGAGGAGACCCGACAGGCGCCGAGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 3510  
 Db 8379 CAATCCCGAGAGGAGACCCGACAGGCGCCGAGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 8438  
 QY 3511 GAGACAGATCCATTCGATTGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 3570  
 Db 8439 GAGACAGATCCATTCGATTGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 8498  
 QY 3571 GCCGTGCTCTTGAAGTACACCGCTTGAAGTACACCGCTTGAAGTACACCGCTTGAAGTACACCG 3630  
 Db 8499 GCCGTGCTCTTGAAGTACACCGCTTGAAGTACACCGCTTGAAGTACACCGCTTGAAGTACACCG 8558  
 QY 3631 TGGAACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3690  
 Db 8559 TGGAACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8618  
 QY 3691 ATTGAGTCAAGAACTTAAGAAATGCTGTTAACTTGAAGTCAAGAACTTGAAGTCAAGAACTTGAAG 3750  
 Db 8619 ATTGAGTCAAGAACTTAAGAAATGCTGTTAACTTGAAGTCAAGAACTTGAAGTCAAGAACTTGAAG 8678  
 QY 3751 TAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3810  
 Db 8679 TAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8738  
 QY 3811 ACATACCTAGAAAGTAAGACAGGAGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3870  
 Db 8739 ACATACCTAGAAAGTAAGACAGGAGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8938  
 QY 3871 TGGTCAAAAAGTACTGTGATTTGATGAGTGGCTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAG 3930  
 Db 8799 TGGTCAAAAAGTACTGTGATTTGATGAGTGGCTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAG 8858  
 QY 3931 CCAGCAGCAGATGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3971  
 Db 8859 CCAGCAGCAGATGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8999

; EARLIER APPLICATION NUMBER: DE 198 56 463  
 ; EARLIER FILING DATE: 1998-11-26  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 15  
 ; LENGTH: 9709  
 ; ORGANISM: Human  
 ; US-09-309-572-15  
 Query Match 43.0%; Score 2676.2; DB 4; Length 9709;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1291 CCATGAGAGTGAAGAGAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350  
 Db 6219 CATGAGAGTGAAGAGAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6278  
 QY 1351 TGTCTCTTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410  
 Db 6279 TGTCTCTTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6338  
 QY 1411 ATGGGATACCTGTGTGAG 1470  
 Db 6339 ATGGGATACCTGTGTGAG 6398  
 QY 1471 CATATGATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530  
 Db 6399 CATATGATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6458  
 QY 1531 ACCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1590  
 Db 6459 ACCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6518  
 QY 1591 TGTGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1650  
 Db 6519 TGTGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6578  
 QY 1651 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1710  
 Db 6579 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6638  
 QY 1711 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1770  
 Db 6639 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6698  
 QY 1771 ATATCAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1830  
 Db 6699 ATATCAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6758  
 QY 1831 ATATGATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1890  
 Db 6759 ATATGATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6818  
 QY 1891 CACAGGCTGTCCAAAGGATCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1950  
 Db 6819 CACAGGCTGTCCAAAGGATCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6878  
 QY 1951 GTTTTGGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2010  
 Db 6879 GTTTTGGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6938  
 QY 2011 TCAGCAGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2070  
 Db 6939 TCAGCAGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6998  
 QY 2071 ATGGCAGTCTAGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130  
 Db 6999 ATGGCAGTCTAGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7058  
 QY 2131 AAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2190  
 Db 7059 AAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7118

RESULT 4  
 US-09-309-572-15  
 ; Sequence 15, Application US/09309572  
 ; Patent No. 6440730  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Heinrich-Pette-Institut  
 ; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV  
 ; FILE REFERENCE: P50489  
 ; CURRENT APPLICATION NUMBER: US/09/309,572  
 ; CURRENT FILING DATE: 1999-05-11

2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGCATTTGTACAATAGGAA 2250  
 |||||||  
 Db ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGCATTTGTACAATAGGAA 7118  
 2251 AAATAGAAATATGAGACACACATTTAATCATAGTATAGAGCAAAATGGAATGCCACTT 2310  
 |||||||  
 Db AAATAGAAATATGAGACACACATTTAATCATAGTATAGAGCAAAATGGAATGCCACTT 7238  
 2311 TAAACAGATATGCTAGCAAAATTTAGAGAACATTTGAAATATATAAACAATATCTTTA 2370  
 |||||||  
 Db TAAACAGATATGCTAGCAAAATTTAGAGAACATTTGAAATATATAAACAATATCTTTA 7298  
 2371 AGCAATCCCTGAGAGGAGGACCGAGAAATTTGTAAGCCACAGTTTAAATGTGGAGGGGAA 2430  
 |||||||  
 Db AGCAATCCCTGAGAGGAGGACCGAGAAATTTGTAAGCCACAGTTTAAATGTGGAGGGGAA 7358  
 2431 TTTTCTACTGTATTAATCAACACACTGTTTAATAGTACTGTTTAAATAGTACTTGGAGTA 2490  
 |||||||  
 Db TTTTCTACTGTATTAATCAACACACTGTTTAATAGTACTGTTTAAATAGTACTTGGAGTA 7418  
 2491 CTGAAGGGTCAAAATACACTGTAAGGAATGACACATACACTCCCATGCGAATATAAC 2550  
 |||||||  
 Db CTGAAGGGTCAAAATACACTGTAAGGAATGACACATACACTCCCATGCGAATATAAC 7478  
 2551 AATTTATTAACATGTCGAGAGAGTGAAGAAAAAGCAATGTATGCCCTCCCATCACTGGAC 2610  
 |||||||  
 Db AATTTATTAACATGTCGAGAGAGTGAAGAAAAAGCAATGTATGCCCTCCCATCACTGGAC 7538  
 2611 AATTTAGATGTTTCATCAATATTTACTGGGCTGCTATTAACAAGAGATGGTGTATTAACA 2670  
 |||||||  
 Db AATTTAGATGTTTCATCAATATTTACTGGGCTGCTATTAACAAGAGATGGTGTATTAACA 7598  
 2671 ACAATGGGTCGAGAGTCTTCAGACCTGGAAGAGGAGATATGAGGACATTTGGAAGAATG 2730  
 |||||||  
 Db ACAATGGGTCGAGAGTCTTCAGACCTGGAAGAGGAGATATGAGGACATTTGGAAGAATG 7658  
 2731 AATTTATTAATATTAATTAAGTAAATTTGAACCATTTGAGATACACCCACCAAGGCA 2790  
 |||||||  
 Db AATTTATTAATATTAATTAAGTAAATTTGAACCATTTGAGATACACCCACCAAGGCA 7718  
 2791 AGAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTGTCTCTGGGT 2850  
 |||||||  
 Db AGAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTGTCTCTGGGT 7778  
 2851 TCTTGGAGCAGAGAGAGACACTATGGGCTGCACGTCAATGACGTGACGGTACAGGCCA 2910  
 |||||||  
 Db TCTTGGAGCAGAGAGAGACACTATGGGCTGCACGTCAATGACGTGACGGTACAGGCCA 7838  
 2911 GACATTTATTTGCTGATATAGTATAGCAGACAGAAATTTGCTGAGGGCTATTTAGGGGC 2970  
 |||||||  
 Db GACATTTATTTGCTGATATAGTATAGCAGACAGAAATTTGCTGAGGGCTATTTAGGGGC 7898  
 2971 AACAGCATCTGTTCACACTCAGCTGTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 3030  
 |||||||  
 Db AACAGCATCTGTTCACACTCAGCTGTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 7958  
 3031 CTGTGGAAGATACTTAAAGATCAACAGCTCTGGGATTTGGGCTGTCTGGAANAAC 3090  
 |||||||  
 Db CTGTGGAAGATACTTAAAGATCAACAGCTCTGGGATTTGGGCTGTCTGGAANAAC 8018  
 3091 TCATTTGGACACCTGCTGCTGGCTTGGAATGCTAGTTGGAGTAATAATCTGGAAGACA 3150  
 |||||||  
 Db TCATTTGGACACCTGCTGCTGGCTTGGAATGCTAGTTGGAGTAATAATCTGGAAGACA 8078  
 8019 TCATTTGGACACCTGCTGCTGGCTTGGAATGCTAGTTGGAGTAATAATCTGGAAGACA 8078  
 3151 TTTGGAATACATGACTGATGAGTGGAGAGAGAAATTAACAATTAACAAGCTTAA 3210  
 |||||||  
 Db TTTGGAATACATGACTGATGAGTGGAGAGAGAAATTAACAATTAACAAGCTTAA 8138  
 8079 TTTGGAATACATGACTGATGAGTGGAGAGAGAAATTAACAATTAACAAGCTTAA 8138  
 3211 TACACTCTTAATTAAGAATGCGAAAAACAGCAAGAAAGATACAGCAATTAATGG 3270  
 |||||||  
 Db TACACTCTTAATTAAGAATGCGAAAAACAGCAAGAAAGATACAGCAATTAATGG 8198  
 8139 TACACTCTTAATTAAGAATGCGAAAAACAGCAAGAAAGATACAGCAATTAATGG 8198

3271 AATTAGTAATATGGGCAAGCTTTGTGAATGGTTTAACATAACAAATGGCTGTGTATA 3330  
 |||||||  
 Db AATTAGTAATATGGGCAAGCTTTGTGAATGGTTTAACATAACAAATGGCTGTGTATA 8258  
 8199 AATTAGTAATATGGGCAAGCTTTGTGAATGGTTTAACATAACAAATGGCTGTGTATA 8258  
 3331 TAAATTTATTAATATGATAGTATAGAGGCTTGTGTATTAAGTAATAGTTTGTGTATC 3390  
 |||||||  
 Db TAAATTTATTAATATGATAGTATAGAGGCTTGTGTATTAAGTAATAGTTTGTGTATC 8318  
 8259 TAAATTTATTAATATGATAGTATAGAGGCTTGTGTATTAAGTAATAGTTTGTGTATC 8318  
 3391 TTTCTATATGATATATGATATAGTATAGGAGATATTCACCTTATTCCTTTAGACCCACTCC 3450  
 |||||||  
 Db TTTCTATATGATATATGATATAGTATAGGAGATATTCACCTTATTCCTTTAGACCCACTCC 8378  
 8319 TTTCTATATGATATATGATATAGTATAGGAGATATTCACCTTATTCCTTTAGACCCACTCC 8378  
 3451 CAATCCGAGGGGAGCCCGACAGGCGCGGAAGAAATGAAGAAGAGGTGAGAGAGACA 3510  
 |||||||  
 Db CAATCCGAGGGGAGCCCGACAGGCGCGGAAGAAATGAAGAAGAGGTGAGAGAGACA 8438  
 8379 CAATCCGAGGGGAGCCCGACAGGCGCGGAAGAAATGAAGAAGAGGTGAGAGAGACA 8438  
 3511 GAGACAGATCCATTCGATTAATGTAAGAGATCTTAAGCACTATTCGAGAGATCTGGGA 3570  
 |||||||  
 Db GAGACAGATCCATTCGATTAATGTAAGAGATCTTAAGCACTATTCGAGAGATCTGGGA 8498  
 8439 GAGACAGATCCATTCGATTAATGTAAGAGATCTTAAGCACTATTCGAGAGATCTGGGA 8498  
 3571 GCGTGGCTCTTCAGTACACACCGCTTGAGAGACTTACTCTGATTTGAACGAGATTCG 3630  
 |||||||  
 Db GCGTGGCTCTTCAGTACACACCGCTTGAGAGACTTACTCTGATTTGAACGAGATTCG 8558  
 8499 GCGTGGCTCTTCAGTACACACCGCTTGAGAGACTTACTCTGATTTGAACGAGATTCG 8558  
 3631 TGGAACTTCTGGGAGCGAGGGGGTGGAAAGCCCTCAAAATATTTGTAATCTCTACAGT 3690  
 |||||||  
 Db TGGAACTTCTGGGAGCGAGGGGGTGGAAAGCCCTCAAAATATTTGTAATCTCTACAGT 8618  
 8559 TGGAACTTCTGGGAGCGAGGGGGTGGAAAGCCCTCAAAATATTTGTAATCTCTACAGT 8618  
 3691 ATTGAGTCAGAACTAAAGAAATAGTGTCTTTAACTGCTCAATGCCACAGCCATAGCAG 3750  
 |||||||  
 Db ATTGAGTCAGAACTAAAGAAATAGTGTCTTTAACTGCTCAATGCCACAGCCATAGCAG 8678  
 8619 ATTGAGTCAGAACTAAAGAAATAGTGTCTTTAACTGCTCAATGCCACAGCCATAGCAG 8678  
 3751 TAGCTGAGGGGACATATAGGCTTATAGAAATTTCAAGACACTTATAGCTATTCGCC 3810  
 |||||||  
 Db TAGCTGAGGGGACATATAGGCTTATAGAAATTTCAAGACACTTATAGCTATTCGCC 8738  
 8679 TAGCTGAGGGGACATATAGGCTTATAGAAATTTCAAGACACTTATAGCTATTCGCC 8738  
 3811 ACATACCTAAGAAATTAAGACAGGCGCTTGAAGAAGATTTGCTATTAAGTGGGTGGCAAG 3870  
 |||||||  
 Db ACATACCTAAGAAATTAAGACAGGCGCTTGAAGAAGATTTGCTATTAAGTGGGTGGCAAG 8798  
 8739 ACATACCTAAGAAATTAAGACAGGCGCTTGAAGAAGATTTGCTATTAAGTGGGTGGCAAG 8798  
 3971 CCAGCAGAGATGGGGTGGAGCACTATCTGAGATCTAGA 3971  
 |||||||  
 Db CCAGCAGAGATGGGGTGGAGCACTATCTGAGATCTAGA 8899  
 8859 CCAGCAGAGATGGGGTGGAGCACTATCTGAGATCTAGA 8899

RESULT 5  
 US-08-646-538-35  
 ; Sequence 35, Application US/08646538  
 ; Patent No. 6027881  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pavlakis, George N.  
 ; APPLICANT: Galtanaris, George A.  
 ; APPLICANT: Steuber, Roland H.  
 ; APPLICANT: Yournakis, John N.  
 ; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIA TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30





Db 7959 CTGTGAAAAGATACCTAAAGATCAACACCTCTGGGGATTTGGGCTGCTGGAAAAC 8018  
 Oy 3091 TCATTTGACACCACTGCTGCTTGGAAATGCTAGTTGGATTAATTAATCTGGAACAGA 3150  
 Db 8019 TCATTTGACACCACTGCTGCTTGGAAATGCTAGTTGGATTAATTAATCTGGAACAGA 8078  
 Oy 3151 TTTGGAAATTAACATGACCTGGATGAGTGGGACAGAAAATTAACATTAACACAGCTTAA 3210  
 Db 8079 TTTGGAAATTAACATGACCTGGATGAGTGGGACAGAAAATTAACATTAACACAGCTTAA 8138  
 Oy 3211 TACACTCTCTTAATTAAGATGCGAAACACAGAAAGAAAGATGACAGATTAATGG 3270  
 Db 8139 TACACTCTCTTAATTAAGATGCGAAACACAGAAAGAAAGATGACAGATTAATGG 8198  
 Oy 3271 AATTGATTAATGAGGCAAGTTGTGGAATGTTTAACTAACATTAAGTGGCTGTGATA 3330  
 Db 8199 AATTGATTAATGAGGCAAGTTGTGGAATGTTTAACTAACATTAAGTGGCTGTGATA 8258  
 Oy 3331 TAAATTTATTCATAATGATAGTAGAGGCTTGTAGTTTAAGATAGTTTGGCTGAC 3390  
 Db 8259 TAAATTTATTCATAATGATAGTAGAGGCTTGTAGTTTAAGATAGTTTGGCTGAC 8318  
 Oy 3391 TTTCTAATGTAATAGATGATGACAGGATATTCACCATTAATGCTTACAGACCACTCC 3450  
 Db 8319 TTTCTAATGTAATAGATGATGACAGGATATTCACCATTAATGCTTACAGACCACTCC 8378  
 Oy 3451 CAATCCGAGGGGACCCGACAGCCCGAAGAAATGAAGAAAGATGAGAGAGAGACA 3510  
 Db 8379 CAATCCGAGGGGACCCGACAGCCCGAAGAAATGAAGAAAGATGAGAGAGAGACA 8438  
 Oy 3511 GAGACAGATCCATTCGATAGTAGAGGCTTGTAGTTTCTGGAGCATCTGGCA 3570  
 Db 8439 GAGACAGATCCATTCGATAGTAGAGGCTTGTAGTTTCTGGAGCATCTGGCA 8498  
 Oy 3571 GCCTTGCTCTTACGCTACCAACCGCTTGAGAGACTTACTTGAATTAACAGAGATTG 3630  
 Db 8499 GCCTTGCTCTTACGCTACCAACCGCTTGAGAGACTTACTTGAATTAACAGAGATTG 8558  
 Oy 3631 TGGAACTTCTGGAGGAGGCGGGTGGGAAAGCCCAATGATGGGAATCTCCACAGT 3690  
 Db 8559 TGGAACTTCTGGAGGAGGCGGGTGGGAAAGCCCAATGATGGGAATCTCCACAGT 8618  
 Oy 3691 ATTGAGTCAGAGATTAAGAAATAGTGTGTTAACTTGTCAATCCACAGCATAGCAG 3750  
 Db 8619 ATTGAGTCAGAGATTAAGAAATAGTGTGTTAACTTGTCAATCCACAGCATAGCAG 8678  
 Oy 3751 TAGCTGAGGGGACAGATAGGCTTATGAAGATTAACAAGCAGCTTATAGAGCTATTGCC 3810  
 Db 8679 TAGCTGAGGGGACAGATAGGCTTATGAAGATTAACAAGCAGCTTATAGAGCTATTGCC 8738  
 Oy 3811 ACATACCTAAGAAATTAAGACAGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 3870  
 Db 8739 ACATACCTAAGAAATTAAGACAGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 8798  
 Oy 3871 TGGTAAAAAAGTAGTGTGTTGATGGCTGCTGTAAGGAAAGATGAGAGCATGAG 3930  
 Db 8799 TGGTAAAAAAGTAGTGTGTTGATGGCTGCTGTAAGGAAAGATGAGAGCATGAG 8858  
 Oy 3931 CCAGCAGCAGATGGGGTGGAGCAGTA 3957  
 Db 8859 CAAGAAATGGCTAGCAAGAGAGAGAA 8885

RESULT 6  
 US-09-503-222-35  
 ; Sequence 35, Application US/09503222  
 ; Patent No. 6265548  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pavlakis, George N.  
 ; APPLICANT: Galanaris, George A.  
 ; APPLICANT: Stauder, Roland H.  
 ; APPLICANT: Vournakis, John N.  
 ; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent

; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/503,222  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/646,538  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Weber, Kenneth A.  
 ; REGISTRATION NUMBER: 31,677  
 ; REFERENCE/DOCKET NUMBER: 015280-249000  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 576-0200  
 ; TELEFAX: (415) 576-0300  
 ; INFORMATION FOR SEQ ID NO: 35:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 15581 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; FEATURE:  
 ; NAME/KEY:  
 ; LOCATION: 1..15581  
 ; OTHER INFORMATION: /note="pNLSG11"  
 ; US-09-503-222-35

Query Match 42.3%; Score 2636.6; DB 3; Length 15581;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 2648; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Oy 1291 CCATGAGCTGAAGAGAGATATCAAGCACTTGTGAGATGGGGTGAATGGGCACCA 1350  
 Db 6219 CAATGAGCTGAAGAGAGATATCAAGCACTTGTGAGATGGGGTGAATGGGCACCA 6278  
 Oy 1351 TGCCTCTGGGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410  
 Db 6279 TGCCTCTGGGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6338  
 Oy 1411 ATGGGGTACCTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG 1470  
 Db 6339 ATGGGGTACCTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG 6398  
 Oy 1471 CATATGATCAGAGGTACATTAATGTTTGGGCCACACATGCTGTGATCCACAGACCCA 1530  
 Db 6399 CATATGATCAGAGGTACATTAATGTTTGGGCCACACATGCTGTGATCCACAGACCCA 6458  
 Oy 1531 ACCCAAGAAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1590  
 Db 6459 ACCCAAGAAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6518  
 Oy 1591 TGGTAGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1650  
 Db 6519 TGGTAGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6578  
 Oy 1651 TAAATTTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATTAATA 1710  
 Db 6579 TAAATTTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATTAATA 6638

1711 CCAATAGTAGAGCGGAGAAATGATATATGAGAAAAAGAGATATAAAAAATGCTCTTTCA 1770  
|||||  
6639 CCAATAGTAGAGCGGAGAAATGATATGAGAAAAAGAGATATAAAAAATGCTCTTTCA 6698  
1771 ATATCAGACAAGAGATAGAGATTAAGTGCAGAAAAAGATATGCACTCTTTATTAACCTTG 1830  
|||||  
6699 ATATCAGACAAGAGATAGAGATTAAGTGCAGAAAAAGATATGCACTCTTTATTAACCTTG 6758  
1831 ATATAGTACCAATAGATATATACAGCTATAGTGTATAGTGTATACCTCAGTCATTA 1890  
|||||  
6759 ATATAGTACCAATAGATATATACAGCTATAGTGTATAGTGTATACCTCAGTCATTA 6818  
1891 CACAGGCTGTCCAAAGGTATCTTTGAGCCAATTTCCCATATTTATTTGCCCCGGCTG 1950  
|||||  
6819 CACAGGCTGTCCAAAGGTATCTTTGAGCCAATTTCCCATATTTATTTGCCCCGGCTG 6878  
1951 GTTTTGCATCTCTAAAAATGTAATATATAGAGTTCAATGGAACGAGGACCATGTACAAATG 2010  
|||||  
6879 GTTTTGCATCTCTAAAAATGTAATATATAGAGTTCAATGGAACGAGGACCATGTACAAATG 6938  
2011 TCAGACAGTACCAATGTACACATGGGAATCAGGCCAGTATGTCAACTCAGCTGTAA 2070  
|||||  
6939 TCAGACAGTACCAATGTACACATGGGAATCAGGCCAGTATGTCAACTCAGCTGTAA 6998  
2071 ATGCACTCTACAGAGAGAGATGTATTAATTAATCTGCAATTTCAAGCAATGCTA 2130  
6999 ATGCACTCTACAGAGAGAGATGTATTAATTAATCTGCAATTTCAAGCAATGCTA 7058  
2131 AACCATATATGTACAGCTGAACACATCTGTATGAATAATTTGTCAAGACCCAAACCA 2190  
7059 AACCATATATGTACAGCTGAACACATCTGTATGAATAATTTGTCAAGACCCAAACCA 7118  
2191 ATCAAGAAAAAGTATCCGTATCCAGAGGGAGCAGGAGAGACATTTGTACAAATAGAA 2250  
7119 ATCAAGAAAAAGTATCCGTATCCAGAGGGAGCAGGAGAGACATTTGTACAAATAGAA 7178  
2251 AAAATGAAATATATGAGACAACACATTTAATCATATAGTAGAGCAAAAAAGGAAATGCTT 2310  
7179 AAAATGAAATATATGAGACAACACATTTAATCATATAGTAGAGCAAAAAAGGAAATGCTT 7238  
2311 TAAACAGATAGTACGCAAAATTAAGAGACAATTTGAAATATATTAACATTAATCTTTA 2370  
7239 TAAACAGATAGTACGCAAAATTAAGAGACAATTTGAAATATATTAACATTAATCTTTA 7298  
2371 AGCAATCTCAGAGAGGGAGCCCAAGAAATTTGAACGACACAGTTTAATTTGTGAGGGAGAT 2430  
7299 AGCAATCTCAGAGAGGGAGCCCAAGAAATTTGAACGACACAGTTTAATTTGTGAGGGAGAT 7358  
2431 TTTTCTACTGTATTAATTAACACAACTGTTTAATAGTACTGTTTAATAGTACTTGAGATA 2490  
|||||  
7359 TTTTCTACTGTATTAATTAACACAACTGTTTAATAGTACTGTTTAATAGTACTTGAGATA 7418  
2491 CTGAAGGGTCAAAATTAACCTGAGAGAGTGAACACATCCCATGCAAGAAATTAAC 2550  
7419 CTGAAGGGTCAAAATTAACCTGAGAGAGTGAACACATCCCATGCAAGAAATTAAC 7478  
2551 AATTTATTAACATGTGCGAGAGAGTAGAGAAAAAGCATGTATGCCCTCCCATAGTGGAC 2610  
7479 AATTTATTAACATGTGCGAGAGAGTAGAGAAAAAGCATGTATGCCCTCCCATAGTGGAC 7538  
2611 AATTTAGTGTTCATCAAAATATTACTGGGCTGTATTAACAAGAGATGGTGTAAATACA 2670  
7539 AATTTAGTGTTCATCAAAATATTACTGGGCTGTATTAACAAGAGATGGTGTAAATACA 7598  
2671 ACAATGGTTCGAGATCTTCAGACCTGAGAGAGGCGATATAGAGGACAATTTGAGAAATG 2730  
7599 ACAATGGTTCGAGATCTTCAGACCTGAGAGAGGCGATATAGAGGACAATTTGAGAAATG 7658  
2731 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2790  
7659 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7718  
2791 AGAGAAAGATGTGTCAGAGAGAAAAAGAGCATGTGGGAATAGAGCTTTGTTCTTGGGT 2850

7729 AGAGAAAGTGTGAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCTTGGGT 7778  
2851 TCTTGGGAGCAGAGAGAGCACTATGGCTGCAGTCAATATACGCTGACGGGTACAGGCA 2910  
7779 TCTTGGGAGCAGAGAGAGCACTATGGCTGCAGTCAATATACGCTGACGGGTACAGGCA 7838  
2911 GACAATTAATGTCTGATATATAGTGAAGCAGCAGCAAAATTTCTGAGGCTATTTAGGGCC 2970  
7839 GACAATTAATGTCTGATATATAGTGAAGCAGCAGCAAAATTTCTGAGGCTATTTAGGGCC 7898  
2971 AACAGCATCTGTGCAACTCAGAGTCTGGGCAATCAACAGCTCCAGGCAAAATCTCTG 3030  
7899 AACAGCATCTGTGCAACTCAGAGTCTGGGCAATCAACAGCTCCAGGCAAAATCTCTG 7958  
3031 CTGTGGAAGATTAATTAAGGATCAACAGCTCTGGGATTTGGGGTCTCTGGAANAC 3090  
7959 CTGTGGAAGATTAATTAAGGATCAACAGCTCTGGGATTTGGGGTCTCTGGAANAC 8018  
3091 TCATTTGCACACAGCTGCTGCTTGAATGCTAGTGGATTAATTAATCTCTGGAACGA 3150  
8019 TCATTTGCACACAGCTGCTGCTTGAATGCTAGTGGATTAATTAATCTCTGGAACGA 8078  
3151 TTTGGAATTAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAACTTAA 3210  
8079 TTTGGAATTAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAACTTAA 8138  
3211 TACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGAAACAAATTAATTTGG 3270  
8139 TACACTCTCTTAATTAAGAAATCGCAAAACAGCAAGAAAGAAACAAATTAATTTGG 8198  
3271 AATTAGATTAATTTGGCAAGTTTGTGAATTTGTTTAACATTAACAAATTTGGCTGTGATA 3330  
8199 AATTAGATTAATTTGGCAAGTTTGTGAATTTGTTTAACATTAACAAATTTGGCTGTGATA 8258  
3331 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3390  
8259 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8318  
3391 TTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3450  
8319 TTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8378  
3451 CAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAGAAAGGTGAGAGAGACA 3510  
8379 CAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAGAAAGGTGAGAGAGACA 8438  
3511 GAGACAGATCCATTCGATTAATGAAGGATTCCTTAGCACTATCTGGGACGATCTGGGA 3570  
8439 GAGACAGATCCATTCGATTAATGAAGGATTCCTTAGCACTATCTGGGACGATCTGGGA 8498  
3571 GCCTGTGCTCTTCAGTACACCGCTTGAAGAGACTTCTCTTGAATTTAAGAGAGATTTG 3630  
8499 GCCTGTGCTCTTCAGTACACCGCTTGAAGAGACTTCTCTTGAATTTAAGAGAGATTTG 8558  
3631 TGAACCTCTTGGAGCGAGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCTACAGT 3690  
8559 TGAACCTCTTGGAGCGAGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCTACAGT 8618  
3691 ATTGGAGTCAGAGAACTAAGAAATATGCTGTTAACTTGAATTCAGCAGCCTATAGCAG 3750  
8619 ATTGGAGTCAGAGAACTAAGAAATATGCTGTTAACTTGAATTCAGCAGCCTATAGCAG 8678  
3751 TAGCTGAGGGAGCAAGATTTGGGTTTATGAAGATTAACAAGCAGCTTATAGACCTATTTGCC 3810  
8679 TAGCTGAGGGAGCAAGATTTGGGTTTATGAAGATTAACAAGCAGCTTATAGACCTATTTGCC 8738  
3811 ACATACCTAGAGAAATTAAGACAGGCTTGAAGGATTTTCTATTAAGATGGGTGGCAAG 3870  
8739 ACATACCTAGAGAAATTAAGACAGGCTTGAAGGATTTTCTATTAAGATGGGTGGCAAG 8798  
3871 TGTCTAAAAAGTACTGTGATTTGATGCTGCTCTGTAAGGGAAGAAATGACGAGCTGAG 3930



Db 7138 TTTGGTCTCTGGAGCAGCAGAGAACTATGGCGGCGAGCGTCAATGACGCTGACGGTAC 7197  
 Oy 2905 AGGCGAGCAATTTATGCTGATATAGTACAGCCAGCAGCAACAATTTGCTGAGCGCTATGG 2964  
 Db 7198 AGGCGAGCAATTTATGCTGATATAGTACAGCAGCAGCAACAATTTGCTGAGCGCTATGG 7257  
 Oy 2965 AGGCGCAACAGCATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGCTCCAGCGAAGAA 3024  
 Db 7258 AGGCGCAACAGCATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGCTCCAGCGAAGAA 7317  
 Oy 3025 TCCTGGCTGTGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTG 3084  
 Db 7318 TCCTGGCTGTGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTG 7377  
 Oy 3085 GAAACATCTTGCACCCCTGCTGCTTGGAAAGCTAGTGGATTAATTAATCTCGG 3144  
 Db 7378 GAAACATCTTGCACCCCTGCTGCTTGGAAAGCTAGTGGATTAATTAATCTCGG 7437  
 Oy 3145 AACAGATTGGAAATACATGACCTGATGAGTGGGAGAGAAATTAACAATTACAGAA 3204  
 Db 7438 AACAGATTGGAAATACATGACCTGATGAGTGGGAGAGAAATTAACAATTACAGAA 7497  
 Oy 3205 GCTTAATACACTCTCTTAATTAAGAAATCCAAACCAAGCAAGAAAAGATGAACAAAT 3264  
 Db 7498 GCTTAATACACTCTCTTAATTAAGAAATCCAAACCAAGCAAGAAAAGATGAACAAAT 7557  
 Oy 3265 TATTGGAATTAGATTAATGGGCAAGTTGGTGAATTTGGTTAACCTAACAAATGGCTGT 3324  
 Db 7558 TATTGGAATTAGATTAATGGGCAAGTTGGTGAATTTGGTTAACCTAACAAATGGCTGT 7617  
 Oy 3325 GGTATATAAATTTATTCATTAATGATAGTAGAGGCTTGGTGAATTAAGTATTTTG 3384  
 Db 7618 GGTATATAAATTTATTCATTAATGATAGTAGAGGCTTGGTGAATTAAGTATTTTG 7677  
 Oy 3385 CTGACTTTCTAATGATAGTATAGTAGGAGAGGATATTCACCATTTATCGTTACAGCC 3444  
 Db 7678 CTGACTTTCTAATGATAGTATAGTAGGAGAGGATATTCACCATTTATCGTTACAGCC 7737  
 Oy 3445 ACCCTCCCAATCCCGAGGGGAGCCGACAGGCGGCGAAGAAATAGAAAGAAAGTGGAGGA 3504  
 Db 7738 ACCCTCCCAATCCCGAGGGGAGCCGACAGGCGGCGAAGAAATAGAAAGAAAGTGGAGGA 7797  
 Oy 3505 GAGACAGAGACAGATCCATTCGATTAGTGAACGAGTCTTACCTATCTGGGAGCATC 3564  
 Db 7798 GAGACAGAGACAGATCCATTCGATTAGTGAACGAGTCTTACCTATCTGGGAGCATC 7857  
 Oy 3565 TCGGAGGCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTTGTATGTAACGA 3624  
 Db 7858 TCGGAGGCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTTGTATGTAACGA 7917  
 Oy 3625 GGATTGTGGAACCTCTGGGAGCGCAGGGGGTGGGAAGCCCTAAATATTTGGGGAATCC 3684  
 Db 7918 GGATTGTGGAACCTCTGGGAGCGCAGGGGGTGGGAAGCCCTAAATATTTGGGGAATCC 7977  
 Oy 3685 TACAGTATTTGAGTCCAGAACCTAAAGATAGTCTGTTAACTTGTCTCAATCCAGAGCA 3744  
 Db 7978 TACAGTATTTGAGTCCAGAACCTAAAGATAGTCTGTTAACTTGTCTCAATCCAGAGCA 8037  
 Oy 3745 TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAATTTCAACAGCCTTATAGAGCTA 3804  
 Db 8038 TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAATTTCAACAGCCTTATAGAGCTA 8097  
 Oy 3805 TTGCGCACATPACCTAGAAAGATTAAGACAGGCTTGGAAAGATTTTGTATTAAGATGGGT 3864  
 Db 8098 TTGCGCACATPACCTAGAAAGATTAAGACAGGCTTGGAAAGATTTTGTATTAAGATGGGT 8157  
 Oy 3865 GGCAGAGTGTCAAAAAGTAGTGTGATTGATGAGCTCTGCTGAAGGAAAGATGAGAGCA 3924  
 Db 8158 GGCAGAGTGTCAAAAAGTAGTGTGATTGATGAGCTCTGCTGAAGGAAAGATGAGAGCA 8217  
 Oy 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGATCTAGA 3971

Db 8218 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCATCTCGAGACTAGA 8264  
 RESULT 8  
 US-08-463-210-4  
 ; Sequence 4, Application US/08463210  
 ; Patent No. 6001977  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHANG, Nancy T.  
 ; APPLICANT: GALLO, Robert C.  
 ; APPLICANT: WONG-STALL, Flossie  
 ; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10154-0053  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463, 210  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/693, 866  
 ; FILING DATE: 23-JAN-1985  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/659, 339  
 ; FILING DATE: 10-OCT-1984  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Serunian, Leslie A.  
 ; REGISTRATION NUMBER: 35,353  
 ; REFERENCE/DOCKET NUMBER: 2026-4193052  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 758-4800  
 ; TELEFAX: (212) 751-6849  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8933 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHEICAL: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: HTLV-III  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; LOCATION: 1..8933  
 ; OTHER INFORMATION: /standard.name="Clone BH10"  
 ; OTHER INFORMATION: /note="Corresponds to nucleotide positions 222 to  
 ; OTHER INFORMATION: 9154 in figure 3 of EP 85307260"  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 113..1648  
 ; OTHER INFORMATION: /product="gag"  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 1408..4452  
 ; OTHER INFORMATION: /product="pol"  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 4367..4975  
 ; OTHER INFORMATION: /product="sor"  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 5560..8148

OTHER INFORMATION: /product= "env"  
us-08-463-210-4

Query Match	41.7%;	Score 2597.8;	DB 3;	Length 8933;
Best Local Similarity	98.2%;	Pred. No. 0;		
Matches 2639;	Conservative 0;	Mismatches 42;	Indels 6;	Gaps 1;
OY 1291	CCATGAGGTGAGAGAGATATGACGACTGTGGAGATGGGGGTGGAATGGGGCACC	1350		
Db				
OY 5579	CAATGAGGTGAGAGAGAAATATCAGCACTTGTGAGATGGGGGTGAGATGGGGCACC	5638		
OY 1351	TGCTCCCTGGGATATGTGATCTGTAGTGTACAGAAAAATTTGGGTCCACCGCTATT	1410		
Db				
OY 5639	TGCTCCCTGGGATGTGTATGTGTGTAGTGTGTACAGAAAAATTTGGGGTCCACCGCTATT	5698		
OY 1411	ATGGGGTACCTGTGTGTGAGAGAGACACCCACTCTATTTTGTGTCATGATGCTAAAG	1470		
Db				
OY 5699	ATGGGGTACCTGTGTGTGAGAGAGACACCCACTCTATTTTGTGTCATGATGCTAAAG	5758		
OY 1471	CATATGATACAGAGGTACATATGTGTGGGCACATGCGGTGTACCCAGACCCCA	1530		
Db				
OY 5759	CATATGATACAGAGGTACATATGTGTGGGCACATGCGGTGTACCCAGACCCCA	5818		
OY 1531	ACCCACAAGAGTATGTGTAATGTGACAGAAAAATTTTACATGTGGAATAATGACA	1590		
Db				
OY 5819	ACCCACAAGAGTATGTGTAATGTGACAGAAAAATTTTACATGTGGAATAATGACA	5878		
OY 1591	TGCTGAGAACATGATGATAGATATATACAGTTTATGGGATCAAAAGCCTAAAGCCTATG	1650		
Db				
OY 5879	TGGTAGAACATGATGATAGATATATACAGTTTATGGGATCAAAAGCCTAAAGCCTATG	5938		
OY 1651	TAAATTTAACCCCACTGTGTGTGTTTAAAGTGACATGATTTGAAGAATGTACTANTA	1710		
Db				
OY 5939	TAAATTTAACCCCACTGTGTGTGTTTAAAGTGACATGATTTGAAGAATGTACTANTA	5998		
OY 1711	CCAAATAGTAGAGCGGGAGAAATGATATGAGAAAAAGAGATAAAAAATGCTCTTTCA	1770		
Db				
OY 5999	CCAAATAGTAGAGCGGGAGAAATGATATGAGAAAAAGAGATAAAAAATGCTCTTTCA	6058		
OY 1771	ATATCAGCAGAGATAGATAGATTAAGGTGAGAAAAAGATATGATGCTTTTATTAACCTG	1830		
Db				
OY 6059	ATATCAGCAGAGATAGATAGATTAAGGTGAGAAAAAGATATGATGCTTTTATTAACCTG	6118		
OY 1831	ATATAGTAGCAATAGATATA-----TACCAGCTATAGTGTGATTAAGTTGTAACACCTGAG	1884		
Db				
OY 6119	ATATAGTAGCAATAGATATA-----TACCAGCTATAGTGTGATTAAGTTGTAACACCTGAG	6178		
OY 1885	TCATTTACACAGGCGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCTATACATTATTTGCC	1944		
Db				
OY 6179	TCATTTACACAGGCGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCTATACATTATTTGCC	6238		
OY 1945	CGGCTGCTTTTGGCATTTCTAAAAATGTATTAATTAAGAGCTTCAATGGAACAGGACCATGTA	2004		
Db				
OY 6239	CGGCTGCTTTTGGCATTTCTAAAAATGTATTAATTAAGAGCTTCAATGGAACAGGACCATGTA	6298		
OY 2005	CAAAATGTAGCAGCAGTACATGTATACATGGAATGAGGCCAGTATATCACTCAACTGCG	2064		
Db				
OY 6299	CAAAATGTAGCAGCAGTACATGTATACATGGAATGAGGCCAGTATATCACTCAACTGCG	6358		
OY 2065	TGTTAAATGGAGCTGACAGAGAAGATGTAGTAATTAAGATCTGCCAATTTTCACAGACA	2124		
Db				
OY 6359	TGTTAAATGGAGCTGACAGAGAAGATGTAGTAATTAAGATCTGCCAATTTTCACAGACA	6418		
OY 2135	ATGCTAAAAACCATTAATAGTACAGCTGAACACATCTGTAGAAATTAATGTACAGACCCA	2184		
Db				
OY 6419	ATGCTAAAAACCATTAATAGTACAGCTGAACACATCTGTAGAAATTAATGTACAGACCCA	6478		
OY 2185	ACAACATATACAAGAAAAAGTATCCGTATCCAGAGAGGACACAGGAGAGCATTTGTTACAA	2244		
Db				
OY 6479	ACAACATATACAAGAAAAAGTATCCGTATCCAGAGAGGACACAGGAGAGCATTTGTTACAA	6538		
OY 2245	TAGAAAAATAGGAATATGAGACAAGCAGATTTTAACTAGTAGAGCAAAATGGAATG	2304		
Db				

Db 6539	TAGAAAAATAGGAATATGAGACAAGCAGCATTTGTACATTAGTAGAGCAAAATGGAATA	6598
OY 2305	CCACTTTAAACAGATAGCTAGCAAAATTAAGGAACAATTTTGGAAATTAATAACAAATA	2364
Db		
OY 6599	ACACTTTAAACAGATAGCTAGCAAAATTAAGGAACAATTTTGGAAATTAATAACAAATA	6658
OY 2365	TCTTTAACCAATCCACAGAGGAGCCAGAAATGTAAAGCAGCTTTTAATTTGTGGAG	2424
Db		
OY 6659	TCTTTAACCAATCCACAGAGGAGCCAGAAATGTAAAGCAGCTTTTAATTTGTGGAG	6718
OY 2425	GGGAATTTTCTACTGTATTTCAACACAATCTTTTAATAGTACTGTTTAATAGTACTT	2484
Db		
OY 6719	GGGAATTTTCTACTGTATTTCAACACAATCTTTTAATAGTACTGTTTAATAGTACTT	6778
OY 2485	GGAGTACGAAGGCTCAAAATTAACACTGAAGAGATGACACAATCACTCCATGACAGA	2544
Db		
OY 6779	GGAGTACGAAGGCTCAAAATTAACACTGAAGAGATGACACAATCACTCCATGACAGA	6838
OY 2545	TAAACAAATTTATAACATGTGGCAGGAAGTGAAGAAAAAGCAATGTATGCCCTCCATCA	2604
Db		
OY 6839	TAAACAAATTTATAACATGTGGCAGGAAGTGAAGAAAAAGCAATGTATGCCCTCCATCA	6898
OY 2605	GTGACAAATTAGATGTTTCATCAAAATTAATTAAGTGGCTCTATTAAACAAGATGGTGTA	2664
Db		
OY 6899	GTGACAAATTAGATGTTTCATCAAAATTAATTAAGTGGCTCTATTAAACAAGATGGTGTA	6958
OY 2665	ATAACAAATTAAGTGGTCCAGATCTTCACACCTGGAGGAGCGATGTAGGGCAATTTGGA	2724
Db		
OY 6959	ATAACAAATTAAGTGGTCCAGATCTTCACACCTGGAGGAGCGATGTAGGGCAATTTGGA	7018
OY 2725	GAAGTAATTTATTAATATTAAGATGAGTAAAAATTTGAACCAATTTAGAGTACACCCACA	2784
Db		
OY 7019	GAAGTAATTTATTAATATTAAGATGAGTAAAAATTTGAACCAATTTAGAGTACACCCACA	7078
OY 2785	AGGCAAGAAGAGAGTGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGAGCTTTGTTC	2844
Db		
OY 7079	AGGCAAGAAGAGAGTGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGAGCTTTGTTC	7138
OY 2845	TTGGTGTCTGGGACACAGAGACCACTATGCGCTGCACGTCATGACGTGAGAGGTAC	2904
Db		
OY 7139	TTGGTGTCTGGGACACAGAGACCACTATGCGCTGCACGTCATGACGTGAGAGGTAC	7198
OY 2905	AGGCGAGCAATTAATGCTGTGATATAGTACAGCAGCAAGCAATTTGCTGAGGCGTATTG	2964
Db		
OY 7199	AGGCGAGCAATTAATGCTGTGATATAGTACAGCAGCAAGCAATTTGCTGAGGCGTATTG	7258
OY 2965	AGGCGCAACAGCATCTGTTGCAACTGCACAGTCTGGGGCATCAACAGCTCCAGCAAGAA	3024
Db		
OY 7259	AGGCGCAACAGCATCTGTTGCAACTGCACAGTCTGGGGCATCAACAGCTCCAGCAAGAA	7318
OY 3025	TCCTGGCTGTGGAAGATATCTTAAGAGATCAACACCTCTGGGGATTTTGGGTTGCTCTG	3084
Db		
OY 7319	TCCTGGCTGTGGAAGATATCTTAAGAGATCAACACCTCTGGGGATTTTGGGTTGCTCTG	7378
OY 3085	GAAAATCTATTTGCAACACTGCTGCTGCTTGGGAATGCTAGTTGGAATTAATCTGCG	3144
Db		
OY 7379	GAAAATCTATTTGCAACACTGCTGCTGCTTGGGAATGCTAGTTGGAATTAATCTGCG	7438
OY 3145	AACAGATTTGGAATTAACATGACCTGATGAGATGGAGTGGAGAGAGAAATTAACAATTAC	3204
Db		
OY 7439	AACAGATTTGGAATTAACATGACCTGATGAGATGGAGTGGAGAGAGAAATTAACAATTAC	7498
OY 3205	GCTTAATACCTCTTAATTTGAAGATGCGCAAAACACAGCAAGAAAGATGAACAAGAT	3264
Db		
OY 7499	GCTTAATACCTCTTAATTTGAAGATGCGCAAAACACAGCAAGAAAGATGAACAAGAT	7558
OY 3265	TATTTGAATTTAGATTAATGAGGCAAGTTTGGGAATTTGGTTTAACATTAACAATTTGCTGT	3324
Db		
OY 7559	TATTTGAATTTAGATTAATGAGGCAAGTTTGGGAATTTGGTTTAACATTAACAATTTGCTGT	7618
OY 3325	GGTATATATAAATTAATTCATTAATGATAGTGAAGGCTTGGTATGTTTAAGAAATGTTTTC	3384
Db		
OY 7619	GGTATATATAAATTAATTCATTAATGATAGTGAAGGCTTGGTATGTTTAAGAAATGTTTTC	7678

QY	3385	CTGTAACCTTCTATAGTAGTAATAGAGTTAGGCGACGGGATATTCACCATTAATCGTTTACAGACC	3444
Db	7679	CTGTAACCTTCTGTAAGTAGTAATAGAGTTAGGCGAGGGATATTAACCATTAATCGTTTACAGACC	7738
QY	3445	ACCTGCCAATCCCGAGGGGACCCGACAGAGCCCGGAAGGAATATGAAGAAGAGGTGAGAGA	3504
Db	7739	ACCTGCCAATCCCGAGGGGACCCGACAGAGCCCGGAAGGAATATGAAGAAGAGGTGAGAGA	7798
QY	3505	GAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTAGCACTTAATCTGGAGCAGTC	3564
Db	7799	GAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTAGCACTTAATCTGGAGCAGTC	7858
QY	3565	TGCGAGAGCCTGTGGCTCTTCACGTACCAACCCGCTTGAGAGACTTAATCTGTGATTTGAACGA	3624
Db	7859	TGCGAGAGCCTGTGGCTCTTCACGTACCAACCCGCTTGAGAGACTTAATCTGTGATTTGAACGA	7918
QY	3625	GGATTGTTGAACCTTCTGGAGCGACAGGGGGTGGAAAGCCCTCAAAATATATGGTGAATCTCC	3684
Db	7919	GGATTGTTGAACCTTCTGGAGCGACAGGGGGTGGAAAGCCCTCAAAATATATGGTGAATCTCC	7978
QY	3685	TACAGTATTTGAGTACGAGACTAAAGAAATAGTGTCTTTAACTTGCTCAATGCCACAGCCA	3744
Db	7979	TACAGTATTTGAGTACGAGAGCTAAAGAAATAGTGTCTTTAACTTGCTCAATGCCACAGCTA	8038
QY	3745	TACAGTATTTGAGGGGAGACAGATAGGGTTATAGAAGTTTCAACAGACACTTAATAGAGCTA	3804
Db	8039	TACAGTATTTGAGGGGAGACAGATAGGGTTATAGAAGTTTCAACAGAGACTTAATAGAGCTA	8098
QY	3805	TTGCGCCACATACCTAGAAAGAAATTAAGACAGGGCTTGAAAAGGATTTTGCCTATTAAGATGGGT	3864
Db	8099	TTGCGCCACATACCTAGAAAGAAATTAAGACAGGGCTTGAAAAGGATTTTGCCTATTAAGATGGGT	8158
QY	3865	GGCAAGTGTCTAAAAAGTAGTGTGATTTGGATGGCCTGCTGTAAAGGAAAGATGAGACGA	3924
Db	8159	GGCAAGTGTCTAAAAAGTAGTGTGATTTGGATGGCCTGCTGTAAAGGAAAGATGAGACGA	8218
QY	3925	GCTGAGCCAGCAGCAGATGGGGTGGAGACAGTATCTCGAGATCTAGA	3971
Db	8219	GCTGAGCCAGCAGCAGATGGGGTGGAGACAGTATCTCGAGACCTAGA	8265
RESULT 9			
US-09-620-958A-3			
: Sequence 3, Application US/09620958A			
: Patent No. 6294338			
: GENERAL INFORMATION:			
: APPLICANT: Nunomura, Kiyotada			
: TITLE OF INVENTION: POLYNOCLEOTIDE AMPLIFICATION METHOD			
: FILE REFERENCE: GPl04-02.0T			
: CURRENT APPLICATION NUMBER: US/09/620.958A			
: CURRENT FILING DATE: 2000-07-21			
: NUMBER OF SEQ ID NOS: 9			
: SOFTWARE: FastSeq for Windows Version 3.0			
: SEQ ID NO 3			
: LENGTH: 8933			
: TYPE: RNA			
: ORGANISM: Human Immunodeficiency Virus			
: FEATURE:			
: NAME/KEY: source			
: LOCATION: (1)...(8933)			
: OTHER INFORMATION: Sequence of transcripts produced from the BH10			
: Patent No. 6294338			
: OTHER INFORMATION: plasmid.			
US-09-620-958A-3			

[illegible][illegible][illegible]







Db	7919	GGANUGUGGAACUUCUGGGACGCGAGGGGGUGGAAAGCCUCCAAAUUADUGUGGAUUCUC	7978
Qy	3685	TACAGTATTGGAGTTCAGGAACCTAAGATAGTCTGTTAACTCTCAATGCCACAGCCA	3744
Db	7979	UACGGUUVUUGGAGGACGAGGACGCUAAAGUAAUUGUGUGUAGCUUUCUCCAAUCCACAGCUA	8038
Qy	3745	TAGCAGTAGGCTGAGAGGGGACAGATATGGGTTATAGAAATATTCAACGACGCTATATAGAGCTA	3804
Db	8039	UAGCAGUAGCGUGAGGGGACACAAUAGCGUUAUAGAAGUAGUACAAGGACCUUAUAGACUA	8098
Qy	3805	TTGCGCCACATPACTAGAAAGAAATTAAGACAGGGCTTGGAAAGGATTTTGCTATAGATGGGT	3864
Db	8099	UUCGCCACAAACCCPAGAAAUAAAGACAGGGCUCUGGAAAGCAUUUUCUAAAPAAUUGGCU	8158
Qy	3865	GGCAAGTGGTCAAAAATAGTGTATGTATGGATGGCCTCTGTAAAGGAAAGATGACAGA	3924
Db	8159	GGCAAGUGGUCAAAAAAGUAGUGUGUGUUGGAAAGGCCUCUUGTAAGGAAAGAAUUGACAGA	8218
Qy	3925	GCTAGGCCACAGCAGAGATGGGGGTGGAGACAGTATCTCGAATCTAGA	3971
Db	8219	GCUTAGCCACAGCAGAUUGGGUGGGGACACACUUCUGAGACCTCAGA	8265

RESULT 11  
US-09-620-958A-9  
; Sequence 9, Application US/09620958A

```

1 Patent No. 6294338
2 GENERAL INFORMATION:
3   APPLICANT: Nunomura, Kiyotada
4   TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
5   FILE REFERENCE: GP104-02.UT
6   CURRENT APPLICATION NUMBER: US/09/620,958A
7   CURRENT FILING DATE: 2000-07-21
8   NUMBER OF SEQ. ID NOS: 9
9   SOFTWARE: FastSeq for Windows Version 3.0
10  SEQ. ID NO. 9
11  LENGTH: 8933
12  TYPE: RNA
13  ORGANISM: Artificial Sequence
14  FEATURE:
15  OTHER INFORMATION: Sequence of the IAC-Bscr pseudo target
16  Patent No. 6294338
17  NAME/KEY: mutation
18  LOCATION: (4140)..(4159)
19  OTHER INFORMATION: Mutated positions: 4140-42, 4145-47, 4152
20  OTHER INFORMATION: 4156-57, 4159
21  IS-09-620-958A-9

```

Query Match	41.78;	Score 2597.8;	DB 3;	Length 8933;
Best Local Similarity	74.38;	Pred. No. 0;		
Matches 1997; Conservative 642;		Mismatches 42;	Indels 6;	Gaps 1;

[illegible]

Db	5879	UGGAGAAACAAGUCACAGACGAAUAAUADCGUUAUGGGGUCGAAACCCGAAAGCCAAUGUGG	5938
Qy	1651	TAAATTTAACCCCACTCTGTGTGTTAGTTTAAAGTCAAGCTGTTTGAAGTAATGATACATAA	1710
Db	5939	UAAAUAUADACCCCACTUCUGUGUUAUGUUUAAAAGGCGACUGAUUUUGAAGAAUAGUACUAAU	5998
Qy	1711	CCAAATAGTAGTAGCGGAGAAATGATAATGAGAAAGAGAGATMAAAACCTGCTCTTTC	1770
Db	5999	CCAAUAGUAGAGCGGGAGAAUAGUAAUUGAGAAAGAGAGAUAAAACUCUGUCUUCUUA	6058
Qy	1771	ATATCAGCACAGCATTAAGAGATTAAGTGCAGAAAGAAATATGCAATTCCTTTATTAACCTG	1830
Db	6059	ADUACACGACAGCAAGCAUAAAGAGUAAAGGUGCAGAAAGAAUUGCGAUUUUUUUAACUUG	6118
Qy	1831	ATATAGTAGCAATATGATTA-----TACGAGCTATAGGTTGATTAAGTGTGAACACCTGAG	1884
Db	6119	AUUAUAUADCCAAUAGUAUAUGAUACUACCGACUUAUGUACACASUUUGUAAACCUCAG	6178
Qy	1885	TCATTTACACAGGCGCTGCCAAAGGTATCTCTTGTAGGCCAAATTCCTCATATTTATGTGCC	1944
Db	6179	UCAUADACACAGGCCUGGCCAAAGGUAAUCUUGAGCGCAAUUCCCAUACAUUAUUGUGGCC	6238
Qy	1945	CGGCTGGTTTTGCGATTCCTTAATATTAATTAAGACGTTCAATGACAGACAGCAGCATGTA	2004
Db	6239	CGGUGUGUUUGCGAUUCCUAAAUGUAUUAAUAAAGACGUCAUUGGACAGCAGCAGUUA	6298
Qy	2005	CAATGTGCAGACAGTACATGTACACATGTGACATGCAAGGCCAGATGATACATCAATCAGCTGC	2064
Db	6299	CAAUUUGCAGCACAGUACAUAUGUACAUUGAAUUGGCCAGAGUAGUACUCAAUCGCAACUCGC	6358
Qy	2065	TGTTAAATGGCAGAGCTGACGAAAGAAAGATATAGTAATGTGATGTCGCAATTCACAGACA	2124
Db	6359	UGUUAUUUGCAGAGUCUGGACGAAAGAGAGAGUAAUUAUAGAUUCUGCAAUUUDCAGCAGAA	6418
Qy	2125	ATGCTAAACCATTAATAGTAGACGCTGAACACATCTGTAGAAATTAATTTGACAGACCA	2184
Db	6419	AUGCUAAAACCAUUAUUGUACAGCUGAACCAUUCUGAGAAUUAUUGUACAAAGACCA	6478
Qy	2185	ACAACATATACAGAAAAAGATATCCGTTTCCAGAGGGGACGAGGAGACGATTTGTTACAA	2244
Db	6479	ACAACAUDACAGAAAAAGAUUCCGUAUCCAGAGAGACGAGGAGAGCAUUDUGUACAA	6538
Qy	2245	TAGAAAAAATAGAAAAATATGAGACACAGCACATTTGTAACTTATAGCAGAAAAATGGAAAG	2304
Db	6539	UAGGAAAAUAUGAAUAUAGAGCAACGACAUUGUAAUUAAGCAAAAAUGGAAU	6598
Qy	2305	CCACTTTAAAACAGATAGTAGCAAAATTAAGAGACAAATTTGGAATTAATTAACACATAA	2364
Db	6599	ACACUUAUAAAACAGAUUGAGUAGCAAAUUAAGAAACAAUUDUGGAAUUAUAAAACAAUUA	6658
Qy	2365	TCTTTAAGCAATCCTCGAGAGGGGAGGCCAGAAATTTGTAACGCACTTTTAATTTGTGGAG	2424
Db	6659	UCUUUUAAGCGAUCCUCGAGGGGAGCCCAAAAUUDUAAACGCAUUDUUUAUUGUGGAG	6718
Qy	2425	GGGAATTTTCTACTGTGAATTTCAACAACAGCTGTTTAAATGACTGTTGTTTAATAGTACTT	2484
Db	6719	GGGAUUUUUUUACUGUAUUCACAACACACUGUUUAUUAUAGUACUGUUUUUAUAGUACU	6778
Qy	2485	GGAGTACTGAAGGCTCAAAATACACTGGAAGAGTAGACACAAATCACACTCCCATCGAGA	2544
Db	6779	GGAGUACUAAAAGGUDCAAAUAAACACUGAAGGAAGUGACAAUACACCCUCCCAUGGAGAA	6838
Qy	2545	TAAACAATTTATTAACATGTGGCAGAGAAAGTAGGAAAAACCAATGTATGCCCCCTCCATCA	2604
Db	6839	UAAAAACAUAUAAACUUDUGGCAGGAAGUAGAAAAACAUAUUGUAGCCUCCUCCACCA	6898
Qy	2605	GTGGAACAATTAGATGTTCACTCAAAATATTACTGGCTGCTATTTAACAACAGATGGTGTA	2664
Db	6899	GUGGACAAAUAGUUGUUCACUCAAUAUUAACAGGGCUGCUAUUUAAACAGAGAGUUGGUA	6958
Qy	2665	ATTAACAACAATGGGTCGAGATCTTGACACTGGAGAGCGCATATGAGGCAATTTGGA	2724

D	6593	AUAGCAACAAUGAUGCCGAGACUUCUCCAGACCUCGAGAGAGAGAUUAGAGGACAAUUGGA	7018
OY	2725	GAAGGAATTTATTAATTAATATAATAGTATTAATAATTTGAACCTTAGAGTAGACACCCACCA	2784
D	7019	GAAGGUAUUUUUAAUAAUAAUAAAGAGUAAAAUUUAAACCAUUVUAGAGUAGACCCACCA	7078
OY	2785	AGCGAAGAGAAAGAGTGTGCAGAGAGAAAAAAGCAGCTGGGAATAGAGACTTTGTTCC	2844
D	7079	AGGCAAAAGAGAGUGUGUGCAGAGAGAAAAAGCAGUGUGGAUUUAGAGACUUUGUCC	7138
OY	2845	TTGGGTTTGGGAGACAGACAGAAAGCACTTAGGGCTGCAGCTCATGAGCGCTGACGGTAC	2904
D	7139	UUGGGUUCUUGGGAGACAGACAGAAAGACUAUUGGGCAGCGCCUCAAUAGCGUCGAGUAC	7198
OY	2905	AGGCCAGACAAATTAATGTCATATAGTCAGACAGAACAAATTTGCTGAGGGCTATTG	2964
D	7199	AGGCCAGACAAUUVUUUGUGUGUAVUGGCGAGCAGACAGAACAAUUVUGUGGGCUNAVG	7258
OY	2965	AGGGCGAACACATCTGTTCGCACTCACAGTCTGGGGCATTAACACGCTCCAGCCAAAGA	3024
D	7259	AGGGCGCAACACACUCUGUGGCACACUCCAGCUCUGGGCAUACACAGCUCGCCAGCAAGAA	7318
OY	3025	TCTTGCGTGTGGAAGAAATPACCTTAAGAGATTCACAGCTCTGGGGATTTGGGGTTCGCTG	3084
D	7319	UCCUGCGUCUGGAAAAAGAACUAAAAGAUCAACACGCUCCUGGGGAUUUGGGGUGUCUCUG	7378
OY	3085	GAAGAACTATTGGCACACACTCGCTGCTGGAAATGCTAGTTGAGATTAATTAATCTCGG	3144
D	7379	GAAGAAUUCUUAUUGGACACACUCGUCGUCUUGGAAUUGCUAGUGAGUAAUUAUUCUCUGG	7438
OY	3145	AACGATTTTGGAAATTAACATGACCTCGGATGGAGTGGGACAGAGAAATTAACATTAACACA	3204
D	7439	AACGAGUUUGGAUAAUACUAGACUCUGGAGUGAGUGGGAGACAGGAAAUUAAACAUUACACAA	7498
OY	3205	GCTTAATPACACTCCTTAATTAAGAAATTCGCAAAACACAGACAGAAAGAAATTAACAAAGAT	3264
D	7499	GCUUAAUUAACUCCUUAAUUGAUAAGAACUCCAAAAACACAGAAAGAAAGAACAAAGAAU	7558
OY	3265	TATTGGAATTAAGATAAATGGGCAAGTTTGTGAAATGGTTTAACATAACAATTTGCGCTGT	3324
D	7559	UAVUGGAUUUAGAUAAAUUGGGCAGAAUUVUUGGAAUUGGUUUAAACAAAUUUGGCGU	7618
OY	3325	GGTATATTAATTAATTCATTAATGATAGTAGAGCGCTTGATGGTTTAAGATAGCTTTTGG	3384
D	7619	GGUU	7678
OY	3385	CTGTAACCTTTCAATGTGTAATGAGTATGAGTGGGAGGGAATATACCACTTAATCGTTGACACC	3444
D	7679	CUGUUCUU	7738
OY	3445	ACCTTCCCAATCCCGAGGGGAGCCGACAGCGCCGAGAAAGAAATAGAAAGAAAGAGTGGAGAGA	3504
D	7739	ACUUCUCCAAUUCCGAGGGGAGCCCGACAGCGCCGAGAAAGAAUUGAAGAAAGAGUUGAGAGA	7798
OY	3505	GAGACAGAGACAGATCCATTCGATTTAGTGAACGAGATCCTTAGCACTTATCTGGAGCGATC	3564
D	7799	GAGACAGAGACAGAUUCCAUUGCAUUDAGUAGACGAGUCCUUVAGCAUUAUUCUGGAGCGAUC	7858
OY	3565	TGCGGAGCGCTGCTCTTCACGCTACACACCGCTTGAGAGACTTACTGTATTTGAACGA	3624
D	7859	UGCGAGCGCUUCGCUUCUCCAGCAACCCGCCUUGAGAGACUUCUUCUUGAUAUUUUUAAACGA	7918
OY	3625	GGATTGTTGAACTTCTGGAGCGCAGGGGGGTGGAGAGCCCTCAAAATATTGGGGAATGCCC	3684
D	7919	GGAUUUGUGAAACUUCUGGGACCGCAGGGGGUGGAAAGCCUCAAUAUUUGGGGGAUUCUCC	7978
OY	3685	TACAGTATTGGAGTCAGAGAACTAAAGATACTGCTGTTTAACCTTGCCTCAATGCCACAGCA	3744
D	7979	UACAGAUUUGAGUACAGGACUCAAAGAAUAGUUGCUUUDAGUUGUCUCUACUAGCCACAGCUA	8038
OY	3745	TACAGATTAAGCTGAGAGGAGACAGATAGGCTTATAGAAGTATTAACAAGACCTTATAGAGCTA	3804
D	8039	UAGCAGUAGUCCGAGGGGACAGAUUAGGUGUUUUAAGAUAUUGUACAAGGAGCUUUAUAGACUA	8098

QY	3805	TTCCGCCACATACCTAGAAAGATTAAGACAGGCGTTGGAAAGATTTTGCTATTAAAGATGGCT	3864
QY	3809	TTCCGCCACATACCTAGAAAGATTAAGACAGGCGTTGGAAAGATTTTGCTATTAAAGATGGCT	3864
Db	8099	TTCCGCCACATACCTAGAAAGATTAAGACAGGCGTTGGAAAGATTTTGCTATTAAAGATGGCT	8158
QY	3865	GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGGAAGAATGAGACGA	3924
QY	8159	GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGGAAGAATGAGACGA	8218
Db	3925	GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGGAAGAATGAGACGA	3971
QY	8219	GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGGAAGAATGAGACGA	8265
Db	3925	GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGGAAGAATGAGACGA	3971

```

RESULT 12
US-09-700-304-1
; Sequence 1, Application US/09700304
; Patent No. 6500623
; GENERAL INFORMATION:
; APPLICANT: GeneCure, LLC
; TITLE OF INVENTION: Replication Defective HIV Vaccines
; FILE REFERENCE: 00714CT05
; CURRENT APPLICATION NUMBER: US/09/700.304
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: PCT/US99/10523
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/085,115
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 9719
; TYPE: DNA
; ORGANISM: Human Immunodeficiency virus type 1
US-09-700-304-1

```

Query Match	41.68;	Score 2593;	DB 4;	Length 9719;
Best Local Similarity	98.1%;	Pred. No. 0;		
Matches 2636; Conservative	0;	Mismatches 45;	Indels 6;	Gaps 1;

QY	1291	CCATGAGAGTGAAGAGAAATATCCAGCACTTGTGAGATGCGGGGTGGAATAATGGGGACCA	1350
Db	6223	CAATTCAGAGTGAAGAGAAATATCCAGCACTTGTGAGATGCGGGGTGGAATAATGGGGACCA	6282
QY	1351	TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAATTGTGGGTCAACCGCTATT	1410
Db	6283	TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAATTGTGGGTCAACAGTCTATT	6342
QY	1411	ATGGGGTACCTGTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG	1470
Db	6343	ATGGGGTACCTGTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG	6402
QY	1471	CATATGATACAGAGGTACATATATGTTTGGGCGACACAGCCCTGTGTAACCCACAGACCCA	1530
Db	6403	CATATGATACAGAGGTACATATATGTTTGGGCGACACAGCCCTGTGTAACCCACAGACCCA	6462
QY	1531	ACCACAGAAGTAGTATTGGTTAAATGTGACAGAAAATTTTACATGTGTGAAAAATGACA	1590
Db	6463	ACCACAGAAGTAGTATTGGTTAAATGTGACAGAAAATTTTACATGTGTGAAAAATGACA	6522
QY	1591	TGTTAGAACCAATGATGATAGATATATATCACTTTATGGGATCAAAGCTTAAAGCATGTG	1650
Db	6523	TGTTAGAACCAATGATGATAGATATATATCACTTTATGGGATCAAAGCTTAAAGCATGTG	6582
QY	1651	TAAATTTAACCCCACTCTGTGTATGTTTAAAGTGCACATGTTTGAGCAATGATACTATA	1710
Db	6583	TAAATTTAACCCCACTCTGTGTATGTTTAAAGTGCACATGTTTGAGCAATGATACTATA	6642
QY	1711	CCAAATAGTAGTAGCGGGAGAAATGATATGAGAAAAGGAGATATAAAAACTGCTCTTCA	1770
Db	6643	CCAAATAGTAGTAGCGGGAGAAATGATATGAGAAAAGGAGATATAAAAACTGCTCTTCA	6702

OY	1777	ATATCAGCACAGCATAAGAGATTAAGGTGCGAAGAAAGATATGCAATCTCTTTATTAACCTG	1830
Db	6703	ATATCAGCACAGCATAAGAGATTAAGGTGCGAAGAAAGATATGCAATCTCTTTATTAACCTG	6762
OY	1831	ATATAGTACCAATAGATA-----TACACACTATAGGTATATAGTTGTATACACCTG	1884
Db	6763	ATATTAATATCCANTGATATATGATCTACACGCTATTAAGTTGACAAGTTGTACACCTG	6822
OY	1885	TCATTTACACAGCCCTGTCCAAAGGTATCTTTGAGCCAAATTCACATTAATGTGCTCC	1944
Db	6823	TCATTTACACAGCCCTGTCCAAAGGTATCTTTGAGCCAAATTCACATTAATGTGCTCC	6882
OY	1945	CGCGTGGTTTTGCCATCTCTAATATGTATTAATTAAGCGTTCAATGAAACAGCACCATGTA	2004
Db	6883	CGCGTGGTTTTGCCATCTCTAATATGTATTAATTAAGCGTTCAATGAAACAGCACCATGTA	6942
OY	2005	CAAAATGTACAGCAGTTCAAATGTACACATGGAATCAGGCGGTGTATTCACATCTGCTC	2064
Db	6943	CAAAATGTACAGCAGTTCAAATGTACACATGGAATCAGGCGGTGTATTCACATCTGCTC	7002
OY	2065	TGTTTAATGTGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAAATTTTCACAGCA	2124
Db	7003	TGTTTAATGTGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAAATTTTCACAGCA	7062
OY	2125	ATGCTTAACCAACATATATAGTACAGCTGAACACATCTGTAGAAATTAATGTATCAAGACCA	2184
Db	7063	ATGCTTAACCAACATATATAGTACAGCTGAACACATCTGTAGAAATTAATGTATCAAGACCA	7122
OY	2185	ACAACATATACAGAAAAAGTATCCGCTATCCAGAGGGGACCGAGGAGACATTTGTTACAA	2244
Db	7123	ACAACATATACAGAAAAAGTATCCGCTATCCAGAGGGGACCGAGGAGACATTTGTTACAA	7182
OY	2245	TAGAAAAAATGTGAAATATGAGACACACATTTTAACATTAATAGACAAATATGGATG	2304
Db	7183	TAGAAAAAATGTGAAATATGAGACACACATTTTAACATTAATAGACAAATATGGATG	7242
OY	2305	CCACTTTAAAACAGATAGCTAGCAAAATTAAGAGAACATTTGGAATATATTAACCAATAA	2364
Db	7243	ACACTTTAAAACAGATAGCTAGCAAAATTAAGAGAACATTTGGAATATATTAACCAATAA	7302
OY	2365	TCCTTTAAGCAATCTCGAGAGGGGACCCAGAAATTTGTAACCCACAGTTTAATGTGGAG	2424
Db	7303	TCCTTTAAGCAATCTCGAGAGGGGACCCAGAAATTTGTAACCCACAGTTTAATGTGGAG	7362
OY	2425	GGGATTTTCTTACTGTATTAATCAACACACAGTTTAATAGACTTGGTTTAATAGTACTT	2484
Db	7363	GGGATTTTCTTACTGTATTAATCAACACACAGTTTAATAGACTTGGTTTAATAGTACTT	7422
OY	2485	GGAGTACTGAAAGGGTCAAAATTAACACTGAAAGGAACTGACACATCACTCCCATGCAAA	2544
Db	7423	GGAGTACTGAAAGGGTCAAAATTAACACTGAAAGGAACTGACACATCACTCCCATGCAAA	7482
OY	2545	TAAANCAATTTATTAACATGTGGGAGGAAGTAGAAGAAAGCAATGTATGCCCTCCATCA	2604
Db	7483	TAAANCAATTTATTAACATGTGGGAGGAAGTAGAAGAAAGCAATGTATGCCCTCCATCA	7542
OY	2605	GTGGAACAAATTAAGATGTTTCATCAATTAATTAATGTGGGCGCTTTTAACAAGAGATGGTGT	2664
Db	7543	GTGGAACAAATTAAGATGTTTCATCAATTAATTAATGTGGGCGCTTTTAACAAGAGATGGTGT	7602
OY	2665	ATTAACCAACATAGGGTCCGAGATCTTTCAGACCTGAGAGGAGGAGATATGAGGACAAATGGA	2724
Db	7603	ATTAACCAACATAGGGTCCGAGATCTTTCAGACCTGAGAGGAGGAGATATGAGGACAAATGGA	7662
OY	2725	GAAATGAAATTTATTAATATTAAGTATGTAATTTGAACCTTTGAGGATGACACCCACCA	2784
Db	7663	GAAATGAAATTTATTAATATTAAGTATGTAATTTGAACCTTTGAGGATGACACCCACCA	7722
OY	2785	AGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGAGAGTGGGAATAGGACCTTTGTTCC	2844
Db	7723	AGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGAGAGTGGGAATAGGACCTTTGTTCC	7782
OY	2845	TTGGGCTTCTGGAGACAGAGAACCTATAGGGCTGCACGCTACATGACGCTGACGGTAC	2904

Db	7783	TTGGGCTTCTGGGAGCAGACAGAAAGCACTAATGGGCGACCTCATATACGTCGACGGTAC	7842
OY	2905	AGGCGACAGCAATATATGCTGATATAGTGCAGCAGCAAGAACAAATTTGCTGAGGCTATTTG	2964
Db	7843	AGGCCAGACAAATTTATGCTGCTGATATAGTGCAGCAGCAGAACAAATTTGCTGAGGCTATTTG	7902
OY	2965	AGGCGCAACAGCAATCTGTTGGCACTCACAGTCTGGGGCATCTMAACAGCTCCAGGCAAGAA	3024
Db	7903	AGGGCAACAGCAATCTGTTGGCACTCACAGTCTGGGGCATCTMAACAGCTCCAGGCAAGAA	7962
OY	3025	TCCTGCGCTGTGGAAAGATACCTTAAAGGATCAACAGCTCTGCGGATTTGGGGTGTGCTGTG	3084
Db	7963	TCCTGCGCTGTGGAAAGATACCTTAAAGGATCAACAGCTCTGCGGATTTGGGGTGTGCTGTG	8022
OY	3085	GAAGAACTAATTTGGACCACTGCTGTGGCTTGGATGCTAGTTGGAGTAATTAATCTGTGG	3144
Db	8023	GAAGAACTAATTTGGACCACTGCTGTGGCTTGGATGCTAGTTGGAGTAATTAATCTGTGG	8082
OY	3145	AACAGATTTGGAAATACATGACCTGCGATGAGTGGAGCGAGAGAAATTAACAAATTTACACAA	3204
Db	8083	AACAGATTTGGAAATACATGACCTGCGATGAGTGGAGCGAGAGAAATTAACAAATTTACACAA	8142
OY	3205	GCTTAATATACACTCTTAATTGAGAAATCGCAAAACAGCAAGAAAGATGACAGAAAT	3264
Db	8143	GCTTAATATACACTCTTAATTGAGAAATCGCAAAACAGCAAGAAAGATGACAGAAAT	8202
OY	3265	TATTTGGAATTTAGTAATTAANTGGGCAAGTTTGTGCAATTTGGTTTAACATTAACAAATTTGGCTGT	3324
Db	8203	TATTTGGAATTTAGTAATTAANTGGGCAAGTTTGTGCAATTTGGTTTAACATTAACAAATTTGGCTGT	8262
OY	3325	GGTATATATTAATTTATTCATATGATGATGAGAGGCTTGTGAGTTTAAGAAATAGTTTGTG	3384
Db	8263	GGTATATATTAATTTATTCATATGATGATGAGAGGCTTGTGAGTTTAAGAAATAGTTTGTG	8322
OY	3385	CTGTACTTTCTAATAGTGAATATAGATTAGGACGGGATATACCAATTTATCGTTTCAGACC	3444
Db	8323	CTGTACTTTCTAATAGTGAATATAGATTAGGACGGGATATACCAATTTATCGTTTCAGACC	8382
OY	3445	ACCTCCCAATCCCGGGGGGACCCGACACAGGCCGGAAGGAATTAAGAAAGAAAGGTGGAGAGA	3504
Db	8383	ACCTCCCAATCCCGGGGGGACCCGACACAGGCCGGAAGGAATTAAGAAAGAAAGGTGGAGAGA	8442
OY	3505	GAGACAGAGACAGATCCATTCCTGATTAAGTAAGGATCCCTTAAGCACTTATCTGGAGCATC	3564
Db	8443	GAGACAGAGACAGATCCATTCCTGATTAAGTAAGGATCCCTTGGCACTTATCTGGAGCATC	8502
OY	3565	TGCGGAGCTGTGCTCTTCAGCTACACCGGCTTGAAGACTTACTCTTGATTTGAACGA	3624
Db	8503	TGCGGAGCTGTGCTCTTCAGCTACACCGGCTTGAAGACTTACTCTTGATTTGAACGA	8562
OY	3625	GGATTGTGGAACTTCTGGGACGCGAGGGGGTGGGAAAGCCCTCAATTAATTTGGTGGAAATCTCC	3684
Db	8563	GGATTGTGGAACTTCTGGGACGCGAGGGGGTGGGAAAGCCCTCAATTAATTTGGTGGAAATCTCC	8622
OY	3685	TACAGTATTTGGAGTCAGGAATTAAGAAATAGTGTGTTAACTTGTCTCAATGCCACACCA	3744
Db	8623	TACAGTATTTGGAGTCAGGAATTAAGAAATAGTGTGTTAACTTGTCTCAATGCCACACCA	8682
OY	3745	TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAATTAATTAACAAGACTTATAGAGCTA	3804
Db	8683	TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAATTAATTAACAAGAGCTTATAGAGCTA	8742
OY	3805	TTTCGCCACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTTGTCTATTAAGATGGGT	3864
Db	8743	TTTCGCCACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTTGTCTATTAAGATGGGT	8802
OY	3865	GGCAAGTGTGTAATAAAGATAGTGTATTTGATGAGGCTCTGTGAAGGAAAGATAGACGA	3924
Db	8803	GGCAAGTGTGTAATAAAGATAGTGTATTTGATGAGGCTCTGTGAAGGAAAGATAGACGA	8862
OY	3925	GCTGAGCCAGCAGAGATGGGGTGGGACAGTATCTCGAGATCTAGA 3971	





QY 3085 GAAACATCTTTGACACACCTGCTGTCCTTGGAAATGCTAGTTGGAGTAATAATCTCTGG 3144  
DB 2066 GAAACATCTTTGACACACCTGCTGTCCTTGGAAATGCTAGTTGGAGTAATAATCTCTGG 2125  
QY 3145 AACAGATTTGGAATACATGACCTGATGAGTGGGACAGAGAAATTAACATTAACACAA 3204  
DB 2126 AACAGATTTGGAATACATGACCTGATGAGTGGGACAGAGAAATTAACATTAACACAA 2185  
QY 3205 GCTTAATACACTCTTAATTAAGAAATGCAAAAACAGCAAGAAAGAAATGACAGAAAT 3264  
DB 2186 GCTTAATACACTCTTAATTAAGAAATGCAAAAACAGCAAGAAAGAAATGACAGAAAT 2245  
QY 3265 TATTTGGAATTAAGTAATGAGGCAAGTTTGGAAATGCTTAACATTAACATTAAGTGGCT 3324  
DB 2246 TATTTGGAATTAAGTAATGAGGCAAGTTTGGAAATGCTTAACATTAACATTAAGTGGCT 2305  
QY 3325 GGTATATTAATTAATTAATGAGTAGAGAGGCTTGGTAAAGTAAAGTAAAGTAAAGT 3384  
DB 2306 GGTATATTAATTAATTAATGAGTAGAGAGGCTTGGTAAAGTAAAGTAAAGTAAAGT 2365  
QY 3385 CTGATCTTTCTAATGTAATGAGTAAAGGAGGATATTCACATTAATCTGTTGACAGACC 3444  
DB 2366 CTGATCTTTCTAATGTAATGAGTAAAGGAGGATATTCACATTAATCTGTTGACAGACC 2425  
QY 3445 ACCATCCAAATCCGAGGAGGAGCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3504  
DB 2426 ACCATCCAAATCCGAGGAGGAGCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2485  
QY 3505 GAGACAGAGACAGATCCATTCATGATAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 3564  
DB 2486 GAGACAGAGACAGATCCATTCATGATAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2545  
QY 3565 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3624  
DB 2546 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2605  
QY 3625 GATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3684  
DB 2606 GATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2665  
QY 3685 TACAGTATTTGAGTACAGAGTAAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCCA 3744  
DB 2666 TACAGTATTTGAGTACAGAGTAAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCCA 2725  
QY 3745 TAGCAGTATTTGAGTACAGAGTAAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCCA 3804  
DB 2726 TAGCAGTATTTGAGTACAGAGTAAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCCA 2785  
QY 3805 TTTGCGCATATCTAGAGAAATAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3864  
DB 2786 TTTGCGCATATCTAGAGAAATAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2845  
QY 3865 GCGAAGTGTCAAAAAGTGTGTGATTTGAGTGGCTGCTTAAGGAAAGATGAGAGCA 3924  
DB 2846 GCGAAGTGTCAAAAAGTGTGTGATTTGAGTGGCTGCTTAAGGAAAGATGAGAGCA 2905  
QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAG 3964  
DB 2906 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCAGTATCTCGAG 2945

RESULT 14  
US-09-318-138-13  
Sequence 13, Application US/09318138  
Patent No. 6531123  
GENERAL INFORMATION:  
APPLICANT: CHANG, Lung-Ji  
TITLE OF INVENTION: LEWNTIVIRAL VECTORS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.  
STREET: 624 Ninth Street, N.W., Suite 300

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/318,138  
FILING DATE: 25-MAY-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/086,635  
FILING DATE: 26-MAY-1998  
APPLICATION NUMBER: US 08/935,312  
FILING DATE: 22-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: CHANG-109A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12479 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-318-138-13  
Query Match 41.1%; Score 2558.4; DB 4; Length 12479;  
Best Local Similarity 99.4%; Pred. NO: 0;  
Matches 2568; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 1291 CCAATGAGTGAAGAGAGATGATACGACCTTGTGAGATGGGGGTGAATGAGGACCA 1350  
DB 5885 CCAATGAGTGAAGAGAGATGATACGACCTTGTGAGATGGGGGTGAATGAGGACCA 5944  
QY 1351 TGCCTCCTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1410  
DB 5945 TGCCTCCTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6004  
QY 1411 ATGGGATACCTGTGTGGAAGAGCAACACACACTCTATTGTCATCAGATGCTAAG 1470  
DB 6005 ATGGGATACCTGTGTGGAAGAGCAACACACACTCTATTGTCATCAGATGCTAAG 6064  
QY 1471 CATATGATACAGAGTACATATGTTGGCCACATGCTGTGTACCCACAGACCCA 1530  
DB 6065 CATATGATACAGAGTACATATGTTGGCCACATGCTGTGTACCCACAGACCCA 6124  
QY 1531 ACCCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1590  
DB 6125 ACCCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6184  
QY 1591 TGTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1650  
DB 6185 TGTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6244  
QY 1651 TAAATTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGATGATGATGAT 1710  
DB 6245 TAAATTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGATGATGATGAT 6304  
QY 1711 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1770  
DB 6305 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6364  
QY 1771 ATATGAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1830

Db 6365 ATATCAGCAGCAAGATTAAGATGAGTGGCAGAAAAAATATGATTCCTTTTAAACCTTG 6424  
Qy 1831 ATATAGTCAATAGATATATACCAAGCTATAGTTGATTAAGTTGTACACCTCAGCATTA 1890  
Db 6425 ATATAGTCAATAGATATATACCAAGCTATAGTTGATTAAGTTGTACACCTCAGCATTA 6484  
Qy 1891 CACAGGCGCTGTCAAAGGATATCTTTGAGCCAAATTCCTATCATATATATGTCGCCGCGTG 1950  
Db 6485 CACAGGCGCTGTCAAAGGATATCTTTGAGCCAAATTCCTATCATATATATGTCGCCGCGTG 6544  
Qy 1951 GTTTTGGCATCTTAAATATGTAATTAAGACGTTCAATGGAACAGGACCATGTACAAATG 2010  
Db 6545 GTTTTGGCATCTTAAATATGTAATTAAGACGTTCAATGGAACAGGACCATGTACAAATG 6604  
Qy 2011 TCACACAGTACATGTACATGTAATCAGGCCAGTGTATGCACTCACTCAGTGGCTTAA 2070  
Db 6605 TCACACAGTACATGTACATGTAATCAGGCCAGTGTATGCACTCACTCAGTGGCTTAA 6664  
Qy 2071 ATGCGACTGTAGCAGAAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCAATGCTA 2130  
Db 6665 ATGCGACTGTAGCAGAAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCAATGCTA 6724  
Qy 2131 AAACCATATATGTACAGTGAACATCTGTAGAATTAATTTGACAAGACCCACAAACA 2190  
Db 6725 AAACCATATATGTACAGTGAACATCTGTAGAATTAATTTGACAAGACCCACAAACA 6784  
Qy 2191 ATACAGAAAAAGATATCCGATTCAGAGGGGAGCAGGAGAGCATTTGTTCAATAGGAA 2250  
Db 6785 ATACAGAAAAAGATATCCGATTCAGAGGGGAGCAGGAGAGCATTTGTTCAATAGGAA 6844  
Qy 2251 AAATAGGAATATGAGACAGACACATTTGTAACTTGTAGAGCAAAATGGAATGCCATT 2310  
Db 6845 AAATAGGAATATGAGACAGACACATTTGTAACTTGTAGAGCAAAATGGAATGCCATT 6904  
Qy 2311 TAAACAGATGTAGCAAAATTAAGAGAACAAATTTGGAATTAATAAACAATATATCTTA 2370  
Db 6905 TAAACAGATGTAGCAAAATTAAGAGAACAAATTTGGAATTAATAAACAATATATCTTA 6964  
Qy 2371 AGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGATTTTAAATTTGAGAGGGAA 2430  
Db 6965 AGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGATTTTAAATTTGAGAGGGAA 7024  
Qy 2431 TTTTCTACTGTAAATCAACACACTGTTTAACTACTGTTTAAATAGTACTTGGAGTA 2490  
Db 7025 TTTTCTACTGTAAATCAACACACTGTTTAACTACTGTTTAAATAGTACTTGGAGTA 7084  
Qy 2491 CTGAAGGCTCAAAATTAACCTGAAGAGTGAACAACTCACTCCCATCGCAATTAAC 2550  
Db 7085 CTGAAGGCTCAAAATTAACCTGAAGAGTGAACAACTCACTCCCATCGCAATTAAC 7144  
Qy 2551 AATTTATTAACATGTGGCAGAGAGTAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC 2610  
Db 7145 AATTTATTAACATGTGGCAGAGAGTAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC 7204  
Qy 2611 AAATTAATGTTCATCAAAATTTACTGGGCTGCTATTAACAAAGATGTGTATTAACA 2670  
Db 7205 AAATTAATGTTCATCAAAATTTACTGGGCTGCTATTAACAAAGATGTGTATTAACA 7264  
Qy 2671 ACAATGGGTCCGAGATCTTCAAGCTGAGAGGCGATATGAGGCAAAATTTGGAGAGTG 2730  
Db 7265 ACAATGGGTCCGAGATCTTCAAGCTGAGAGGCGATATGAGGCAAAATTTGGAGAGTG 7324  
Qy 2731 AATTTATTAATTAATTAAGATTAATAAATTTGAACCATTAAGAGTGTAGCAACCAAGGCAA 2790  
Db 7325 AATTTATTAATTAATTAAGATTAATAAATTTGAACCATTAAGAGTGTAGCAACCAAGGCAA 7384  
Qy 2791 AGAGAAAGTGTGTCAGAGAAAAAGAGCAGTGGGAATGAGAGCTTTGTTCTTGGGT 2850  
Db 7385 AGAGAAAGTGTGTCAGAGAAAAAGAGCAGTGGGAATGAGAGCTTTGTTCTTGGGT 7444  
Qy 2851 TCTTGGAGAGCAGGAAGCACTATGGGCGACGCTCAATGACGCTGACGTTACAGGCCA 2910  
Db 7445 TCTTGGAGAGCAGGAAGCACTATGGGCGACGCTCAATGACGCTGACGTTACAGGCCA 7504

Qy 2911 GACAAATATGTCTGATATAGTGCAGCAGCAAGCAATTTGTCGAGGCGCTTATGAGCGC 2970  
Db 7505 GACAAATATGTCTGATATAGTGCAGCAGCAAGCAATTTGTCGAGGCGCTTATGAGCGC 7564  
Qy 2971 AACAGATCTGTTGCAACCTCAGACCTGTGGGGCATCAACAGCTCCAGGCAAGATCTGCG 3030  
Db 7565 AACAGATCTGTTGCAACCTCAGACCTGTGGGGCATCAACAGCTCCAGGCAAGATCTGCG 7624  
Qy 3031 CTGTGGAAGATTAACCTTAAGATTAACAGCTCTGCGGGATTTGGGGTTGCTGTGGAAC 3090  
Db 7625 CTGTGGAAGATTAACCTTAAGATTAACAGCTCTGCGGGATTTGGGGTTGCTGTGGAAC 3090  
Qy 3091 TCATTTGACACCACTGCTGCTTGAATGCTAGTGTAGTAATTAATCTGTGGAACA 3150  
Db 7685 TCATTTGACACCACTGCTGCTTGAATGCTAGTGTAGTAATTAATCTGTGGAACA 7744  
Qy 3151 TTTGGAATTAACATGACCTGTGAGTGGGACAGAAATTTAACAAATTAACAAGCTTAA 3210  
Db 7745 TTTGGAATTAACATGACCTGTGAGTGGGACAGAAATTTAACAAATTAACAAGCTTAA 7804  
Qy 3211 TACACCTCTTAATTAAGAAATCCGAAAAACAGCAAGAAAGAAATGAACAAAGATTAATGG 3270  
Db 7805 TACACCTCTTAATTAAGAAATCCGAAAAACAGCAAGAAAGAAATGAACAAAGATTAATGG 7864  
Qy 3271 AATTAAGATTAATGAGGCAAGTTTGTGAATTTGTTTAACTAATTAACAAATTTGCTGTATA 3330  
Db 7865 AATTAAGATTAATGAGGCAAGTTTGTGAATTTGTTTAACTAATTAACAAATTTGCTGTATA 7924  
Qy 3331 TAAATTTATTCATATATATATATAGTAGAGGCTTGTGATTAAGAAATGTTTTGCTGTAC 3390  
Db 7925 TAAATTTATTCATATATATATATATAGTAGAGGCTTGTGATTAAGAAATGTTTTGCTGTAC 7984  
Qy 3391 TTTCTATAGGAATAGAGTTAGAGGAGATTAATCAGATTAATGTTTCAACACCACTCC 3450  
Db 7985 TTTCTATAGGAATAGAGTTAGAGGAGATTAATCAGATTAATGTTTCAACACCACTCC 8044  
Qy 3451 CAATCCCGAGGGGAGCCGACAGGCGCCGAAGAAATGAAGAAAGAGTGTGAGAGAGACA 3510  
Db 8045 CAATCCCGAGGGGAGCCGACAGGCGCCGAAGAAATGAAGAAAGAGTGTGAGAGAGACA 8104  
Qy 3511 GAGACAGATCCATTCGATTAAGTAGAGAGATCTTAAGCATTATCTGGACGATCTCGCGA 3570  
Db 8105 GAGACAGATCCATTCGATTAAGTAGAGAGATCTTAAGCATTATCTGGACGATCTCGCGA 8164  
Qy 3571 GCCGTGCTCTTCAAGTACACCGCTTGAAGAGCTTACTGATTTGAACGAGATTTG 3630  
Db 8165 GCCGTGCTCTTCAAGTACACCGCTTGAAGAGCTTACTGATTTGAACGAGATTTG 8224  
Qy 3631 TGAACCTTCTGGGACGAGGGGTTGGGAACCCCTCAATATTTGGTGAATCTCTACAGT 3690  
Db 8225 TGAACCTTCTGGGACGAGGGGTTGGGAACCCCTCAATATTTGGTGAATCTCTACAGT 8284  
Qy 3691 ATTGGAATCAGGAATTAAGAAATTAAGTCTTTAACTTGTCAATGTCACAGCATATAGAG 3750  
Db 8285 ATTGGAATCAGGAATTAAGAAATTAAGTCTTTAACTTGTCAATGTCACAGCATATAGAG 8344  
Qy 3751 TAGCTGAGGGACAGATAGGTTATAGAAATTAATTAACAGCAGCTTATAGAGCTTATTCGCC 3810  
Db 8345 TAGCTGAGGGACAGATAGGTTATAGAAATTAATTAACAGCAGCTTATAGAGCTTATTCGCC 8404  
Qy 3811 ACATACCTAGAAAGATTAAGACAGAGGCTTGAAGAGATTTTGCATTAAGATGAGGTGCAAG 3870  
Db 8405 ACATACCTAGAAAGATTAAGACAGAGGCTTGAAGAGATTTTGCATTAAGATGAGGTGCAAG 3870  
Qy 3871 TGGT 3874  
Db 8465 TGGT 8468

RESULT 15  
US-08-935-312-13  
; Sequence 13, Application US/08935312

Patent No. 6207455  
: GENERAL INFORMATION:  
: APPLICANT: CHANG, Lung-Ji  
: TITLE OF INVENTION: LENTIVIRAL VECTORS  
: NUMBER OF SEQUENCES: 20  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: BROWDY AND NEIMARK, P. L. L. C.  
: STREET: 624 Ninth Street, N.W.  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20001  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentln Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/935,312  
: FILING DATE: 22-SEP-1997  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: COOPER, Iver P.  
: REGISTRATION NUMBER: 28,005  
: REFERENCE/DOCKET NUMBER: CHANG-112  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 202-628-5197  
: TELEFAX: 202-737-3528  
: INFORMATION FOR SEQ ID NO: 13:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 12494 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: other nucleic acid  
: DESCRIPTION: /desc = "DNA"  
: US-08-935-312-13

Query Match 41.1%; Score 2558.4; DB 3; Length 12494;  
Best Local Similarity 99.4%; Pred. NO. 0; Mismatches 16; Indels 0; Gaps 0;  
Matches 2568; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 1291 CCATGAGAGTGAAGAGAGATATCAGCACTGTGTGAGATGGGGTGGAAATGGGGCAACA 1350  
DB 5890 CAATGAGAGTGAAGAGAGATATCAGCACTGTGTGAGATGGGGTGGAAATGGGGCAACA 5949  
QY 1351 TGTCTCTTGGGATATGTGATGATCTGTAGTGTACAGAAAATTTGTGGGTACCGCTTATT 1410  
DB 5950 TGTCTCTTGGGATATGTGATGATCTGTAGTGTACAGAAAATTTGTGGGTACCGCTTATT 6009  
QY 1411 ATGGGGTACCTGTGTGAGAGAACCAACCACTCTATTTTGTGATCAATGCTTAAG 1470  
DB 6010 ATGGGGTACCTGTGTGAGAGAACCAACCACTCTATTTTGTGATCAATGCTTAAG 6069  
QY 1471 CATATGATACAGAGTACATATATTTGGGGCCACACATGCTGTATACCCACAGACCCA 1530  
DB 6070 CATATGATACAGAGTACATATATTTGGGGCCACACATGCTGTATACCCACAGACCCA 6129  
QY 1531 ACCCACAAGAGTATGTTGTAATGTGACAGAAAATTTTAAACATGTGAAAAATGACA 1590  
DB 6130 ACCCACAAGAGTATGTTGTAATGTGACAGAAAATTTTAAACATGTGAAAAATGACA 6189  
QY 1591 TGTGTAGAACAGATGAGATATATCACTTATGAGTCAAAAGCCTTAAGGCATGTG 1650  
DB 6190 TGTGTAGAACAGATGAGATATATCACTTATGAGTCAAAAGCCTTAAGGCATGTG 6249  
QY 1651 TAAATTTAAACCCACTGTGTGTTAGTTTAAAGTCACTGATTTGAAGATGATCTATA 1710  
DB 6250 TAAATTTAAACCCACTGTGTGTTAGTTTAAAGTCACTGATTTGAAGATGATCTATA 6309  
QY 1711 CCAATAGTAGTAGCGGGAGAAATGATATGAGAAAAGAGATAAAAATGCTCTTTCA 1770  
DB 6310 CCAATAGTAGTAGCGGGAGAAATGATATGAGAAAAGAGATAAAAATGCTCTTTCA 6369

OY 1771 ATATCAGACAAGCATAGATATAGTGCAGAAAAGATATGATCTTTTATTAACCTTG 1830  
DB 6370 ATATCAGACAAGCATAGATATAGTGCAGAAAAGATATGATCTTTTATTAACCTTG 6429  
OY 1831 ATATAGTACCAATAGATATATACAGCTATAGTTGATTAAGTTGTAAACCTCACTCATTA 1890  
DB 6430 ATATAGTACCAATAGATATATACAGCTATAGTTGATTAAGTTGTAAACCTCACTCATTA 6489  
OY 1891 CACAGGCTGTCCAAAGGTATCTTTGAGCCATTTCCCATATATTATTTGCCCCGGCTG 1950  
DB 6490 CACAGGCTGTCCAAAGGTATCTTTGAGCCATTTCCCATATATTATTTGCCCCGGCTG 6549  
OY 1951 GTTTGGATCTTAAATGTATATATTAAGCTTCAATGTGAACGAGACATGTACAAATG 2010  
DB 6550 GTTTGGATCTTAAATGTATATATTAAGCTTCAATGTGAACGAGACATGTACAAATG 6609  
OY 2011 TCAGCAGATCAATGTACACATGTGAATCAGGCCAGTATGTATCACTCACTGCTTTAA 2070  
DB 6610 TCAGCAGATCAATGTACACATGTGAATCAGGCCAGTATGTATCACTCACTGCTTTAA 6669  
OY 2071 ATGCAGCTTAGCAGAGAAATGTATATTTAGATCTGCCAATTTTACAGACAATGCTA 2130  
DB 6670 ATGCAGCTTAGCAGAGAAATGTATATTTAGATCTGCCAATTTTACAGACAATGCTA 6729  
OY 2131 AAACCATATATGTACAGCTGAACACATCTGTAGAATTAATGTACAAAGACCAACACA 2190  
DB 6730 AAACCATATATGTACAGCTGAACACATCTGTAGAATTAATGTACAAAGACCAACACA 6789  
OY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGAGCAGGAGAGCATTTGTTCAATAGGAA 2250  
DB 6790 ATACAGAAAAAGTATCCGTATCCAGAGGGGAGCAGGAGAGCATTTGTTCAATAGGAA 6849  
OY 2251 AAATAGCAATATATGAGACACACATTTTAACTTAGTAGAGCAAAATGGAATGCACTT 2310  
DB 6850 AAATAGCAATATATGAGACACACATTTTAACTTAGTAGAGCAAAATGGAATGCACTT 6909  
OY 2311 TAAACAGATAGTACGCAAAATTAAGAGAACATTTGGAATTAATTAATCAATATCTTTA 2370  
DB 6910 TAAACAGATAGTACGCAAAATTAAGAGAACATTTGGAATTAATTAATCAATATCTTTA 6969  
OY 2371 AGCAATCTCAGAGAGGGAGCCAGAAATTTGTAGCCACAGTTTAAATTTGTGAGGGGAAT 2430  
DB 6970 AGCAATCTCAGAGAGGGAGCCAGAAATTTGTAGCCACAGTTTAAATTTGTGAGGGGAAT 7029  
OY 2431 TTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTGGAATA 2490  
DB 7030 TTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTGGAATA 7089  
OY 2491 CTGAAGGCTCAATTAACCTGAAGAGTGAACACATCACTCCCATGCAAGATTAANAC 2550  
DB 7090 CTGAAGGCTCAATTAACCTGAAGAGTGAACACATCACTCCCATGCAAGATTAANAC 7149  
OY 2551 AATTTATTAACATGTGCGAGAGAGTAGGAAAAGCAATGTATGCCCTCCCATAGTGGAC 2610  
DB 7150 AATTTATTAACATGTGCGAGAGAGTAGGAAAAGCAATGTATGCCCTCCCATAGTGGAC 7209  
OY 2611 AAATTTAGTGTTCATCAATTTTACTGGGCTGCTATTTAACAAGATGTTGTATTAACA 2670  
DB 7210 AAATTTAGTGTTCATCAATTTTACTGGGCTGCTATTTAACAAGATGTTGTATTAACA 7269  
OY 2671 ACAATGGTCCGAGATCTTACAGCTGGAAGAGGGAATATAGGAGCAATTTGGAGAAATG 2730  
DB 7270 ACAATGGTCCGAGATCTTACAGCTGGAAGAGGGAATATAGGAGCAATTTGGAGAAATG 7329  
OY 2731 AATTTATTAATTAATTAAGTATTAATTTGAACATTTAGAGATAGCACCCACAAGGCAA 2790  
DB 7330 AATTTATTAATTAATTAAGTATTAATTTGAACATTTAGAGATAGCACCCACAAGGCAA 7389  
OY 2791 AGAGAAAGTGTGTGAGAGAGAAAAAGAGAGAGTGGGAATAGAGCTTTGTTCTTGCGT 2850  
DB 7390 AGAGAAAGTGTGTGAGAGAGAAAAAGAGAGAGTGGGAATAGAGCTTTGTTCTTGCGT 7449

Search completed: September 17, 2003, 17:30:45  
Job time : 337 secs

```
QY 2851 TCTTGGAGCAGCAGGAAGCACTATGGGCTGCAGTCAATGACGCTGACGGGTACAGGCCA 2910
Db 7450 TCTTGGAGCAGCAGGAAGCACTATGGGCTGCAGTCAATGACGCTGACGGGTACAGGCCA 7509
QY 2911 GACAAATTATTTCTGATATAGTCAGCAGCAGGAACAATTTGCTGAGGGCTATTGAGCGCC 2970
Db 7510 GACAAATTATTTCTGATATAGTCAGCAGCAGGAACAATTTGCTGAGGGCTATTGAGCGCC 7569
QY 2971 AACAGCACTCTGTCAACTCAGACTCTGGGCTCAACACACCTCCAGGCAAGAAATCTCTGG 3030
Db 7570 AACAGCACTCTGTCAACTCAGACTCTGGGCTCAACACACCTCCAGGCAAGAAATCTCTGG 7629
QY 3031 CTGTGAAAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTTGCTCTGAAAAAC 3090
Db 7630 CTGTGAAAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTTGCTCTGAAAAAC 7689
QY 3091 TCATTTGCACACGCTGTGCTGTGGAATGCTATGTGAGTAATTAATCTCTGGAACAGA 3150
Db 7690 TCATTTGCACACGCTGTGCTGTGGAATGCTATGTGAGTAATTAATCTCTGGAACAGA 7749
QY 3151 TTTGGAATTAACATGACCTGATGAGTGAGCAGAGAAATTAACAAATTACACAAGCTTAA 3210
Db 7750 TTTGGAATTAACATGACCTGATGAGTGAGCAGAGAAATTAACAAATTACACAAGCTTAA 7809
QY 3211 TACACTCCTTAATTGAAGATCGCAAAACGACAGAAAAGAAATGAACAGAAATTAATGG 3270
Db 7810 TACACTCCTTAATTGAAGATCGCAAAACGACAGAAAAGAAATGAACAGAAATTAATGG 7869
QY 3271 AATTAGATTAATGGGCAAGTTTGGAAATTTGGTTAACAATAATTTGGCTGTGATTA 3330
Db 7870 AATTAGATTAATGGGCAAGTTTGGAAATTTGGTTAACAATAATTTGGCTGTGATTA 7929
QY 3331 TAAATTTATTCATATGATAGTAGAGGCTTGTAGTTTAAGAAATGTTTGTCTGTAC 3390
Db 7930 TAAATTTATTCATATGATAGTAGAGGCTTGTAGTTTAAGAAATGTTTGTCTGTAC 7989
QY 3391 TTTCTATAGTGAATAGATTAGCGAGGAAATTCACACTTAATCGTTTCAGACCACTCC 3450
Db 7990 TTTCTATAGTGAATAGATTAGCGAGGAAATTCACACTTAATCGTTTCAGACCACTCC 8049
QY 3451 CAATCCGAGGGGACCCGACAGCGCCGAAAGGAATAGAAAGAGTGAGAGAGACA 3510
Db 8050 CAATCCGAGGGGACCCGACAGCGCCGAAAGGAATAGAAAGAGTGAGAGAGACA 8109
QY 3511 GAGACAGATCCATTGATTAAGTGAACGATCCTTAGCACTTAATCTGGGACGATCTGCGGA 3570
Db 8110 GAGACAGATCCATTGATTAAGTGAACGATCCTTAGCACTTAATCTGGGACGATCTGCGGA 8169
QY 3571 GCCGTGCTCTTACGCTACACCGCTTGAGAGACTTACTGTTGATTTGAAGAGATTG 3630
Db 8170 GCCGTGCTCTTACGCTACACCGCTTGAGAGACTTACTGTTGATTTGAAGAGATTG 8229
QY 3631 TGSAACTTCTGGGACGAGGGGGTGGGAAGCCCTAAATATTTGGGAATCTCTACAGT 3690
Db 8230 TGSAACTTCTGGGACGAGGGGGTGGGAAGCCCTAAATATTTGGGAATCTCTACAGT 8289
QY 3691 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGCATGCAATGCACACAGCATATAGCAG 3750
Db 8290 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGCATGCAATGCACACAGCATATAGCAG 8349
QY 3751 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTACAGACGCTTATAGAGCTATTGCCC 3810
Db 8350 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTACAGACGCTTATAGAGCTATTGCCC 8409
QY 3811 ACATACCTAGAAGAAATAGACAGGGCTTGAAGAGATTTGCTTAAGATGGGTGCAAG 3870
Db 8410 ACATACCTAGAAGAAATAGACAGGGCTTGAAGAGATTTGCTTAAGATGGGTGCAAG 8469
QY 3871 TGGT 3874
Db 8470 TGGT 8473
```







OY	3685	TACAGTATTGGATCAGACACTAAAGATAGTGTGTTAACTGCTCAATGGCCAGCCA	3744
Db	8802	TACAGTATTGGATCAGACACTAAAGATAGTGTGTTAACTGCTCAATGGCCAGCCA	8861
OY	3745	TAGCAGTAGTGAGGGGACAGATAGGCTTATAGAAATTAACAAGCAGCTTATAGAGCTA	3804
Db	8862	TAGCAGTAGTGAGGGGACAGATAGGCTTATAGAAATTAACAAGCAGCTTATAGAGCTA	8921
OY	3805	TTCCGCACATACCTTGAAGATTAAGACAGGGCTTGGAAAGAGTTTGGCTATTAAGTGGCT	3864
Db	8922	TTCCGCACATACCTTGAAGATTAAGACAGGGCTTGGAAAGAGTTTGGCTATTAAGTGGCT	8981
OY	3865	GGCAAGTGGTAAAAAGTAGTGTGATTTGATGGCCCTGCTTAAGGGAAGATGAGACA	3924
Db	8982	GGCAAGTGGTAAAAAGTAGTGTGATTTGATGGCCCTGCTTAAGGGAAGATGAGACA	9041
OY	3925	GCTGAGCCAGACAGATGGGGTGGAGCAGTATCTGA	3963
Db	9042	GCTGAGCCAGACAGATGGGGTGGAGCAGCATTCTGAGTATATTAAAGCAGAACTTGT	9101
OY	3964	-----	3963
Db	9102	TTATTTGACGCTTAATAGTGTACAAATAAGCAATAGATCACAATAATTTACAAATAAG	9161
OY	3964	-----	3963
Db	9162	CATTTTTTCACTGATTCATAGTTGTGTGTGCCAACTCATCATATCTTATTCATG	9221
OY	3964	-----	3963
Db	9222	TCTGTGCACTGTAGAGCTAAGCAAGTATGCAAAATGCAGTGCTTCGCATTTGGAAA	9281
OY	3964	-----	3963
Db	9282	TCCCCAATCTGGTGGCCAAAGAGGTTCAATAATACAGTAATGGAACCTCTTATGTCAAAAGT	9341
OY	3964	-----	3982
Db	9342	GAGTACTGTTTAACAGCGTAGATTGGAAACAGCCTTCAGAGATGCTTTGGACGTGGCAGTTGT	9401
OY	3983	TTTCATCCAGACATGATTAAGTACATTGATGATTTGGACAAACCA	4028
Db	9402	TCCACTGCAMAGACATCACAAACTTCATGATGATGTTGTACACACAGGTCCCTGMAAGGC	9461
OY	4029	-----	4028
Db	9462	AACCTACTTCACTGCAAAAGTGAACATAAAGGAATGGTTTCATCATGCGCCGTGAT	9521
OY	4029	-----	4028
Db	9522	GCCACTGACATAGAAAACGAATGACAAGAGCGATGAAGTATTTGGTGGCAGATGG	9581
OY	4029	-----	4068
Db	9582	TCAGAGTCGAGAGAGTGTCCACCCAGCTTGTGGTGGCCCATATCTTCAACCACTAGAA	9641
OY	4069	TTTGATGCTAFTGTGTTATTTGTTAACCATTATAAGTCGCAATAAACAAGTTAAACAATA	4128
Db	9642	TAAACACTTACGTATCATTAATTAATTAAGCATATCAGTACCAATTAATTCAGAAATCAGCAAA	9701
OY	4129	CAATTCATTCATTTTATGTTTCAAGTT	4156
Db	9702	CAGTCATGTTTATATATATCTTGTGCTTAATTCATGACAGATATTTATTTTATAGT	9761
OY	4157	-----	4159
Db	9762	ACAGTTTACTGTTCTCTCCACCAATGCTCTCTGCAATTAAGGCTGCAAGTGTGACTAG	9821
OY	4160	GGGGAGCTGGGGAGGTTTAAAGCAAGTAAACCTCTACAATGTGGATGGCTGAT	4219
Db	9822	CAGAAAGTTCAAGTAACTCTTGAAGACTGCAAAATTCATTTAGCTGGTCAAAAGACTTGA	9881

---

OY	4220	TATGATCT	4228
Db	9882	TTTGTCTATTTAATTTTCCCTTTTCTCTCATATAACTTTTAAGCCACTGAGATGAAT	9941
OY	4229	-----	4228
Db	9942	TCTGGAACATAAGTAGTAATAATACCTTCCTGACTTGTAGAAAATGTGAATACTAGATGC	10001
OY	4229	-----	4281
Db	10002	CACCTAGTCTCTGTGTAGTACTGTATTAACACAGCAATTCGGCTATATATGCTTACTACT	10061
OY	4282	G-----	4322
Db	10082	GTTTCCTGTGTGAATTTTATTCGCTCATCAATTCACACAACATACAGAGCCGGAAGAT	10121
OY	4333	A-----	4333
Db	10122	AAAGTAAAGCCTGGGGTGCTTAATGATGAGTAACTCACTCAATTAATTTGGTGGCTGC	10181
OY	4334	AGGGCGCTCAAGCGGCTTTGGCGGGTGTGGGGGGCAGCCATGACCCAGTCAGTAGCG	4393
Db	10182	ACTGCCCTTTCCAGTCGGGAAACCTGTGCGCAGCTGCATTAAATGAATGCGCAACG	10241
OY	4394	ATAGCGAGTGTATCTGGCTTAA	4417
Db	10242	CGCGGGAGAGCGGTTTGCATTTGGGCGCTTCGCTTCCTGCTCAGTACGCTGCT	10301
OY	4418	-----	4417
Db	10302	GCCTCGGCTGTGGCTGCGGCGAGCGGATACACTCACTCAAAAGCGGTAATACGGTT	10361
OY	4418	-----	4462
Db	10362	ATCCACAGATACAGGGGATTAACGACGAAGAAACATGTGAGCAAAAAGCCAGCAAAAAGCC	10421
OY	4463	-----	4508
Db	10422	CAGGAAGCTAAAAAGCGCGGTTGCTGCGGTTTTCATAGGCTCGCGCCCTCGACGA	10481
OY	4509	GCATCACAAAAATGACGCTCAAGTCACAGATGGCGAAAACCCGACAGACATTAAGAATA	4568
Db	10482	GCATCACAAAAATGACGCTCAAGTCACAGATGGCGAAAACCCGACAGACATTAAGAATA	10541
OY	4569	CGAGGCTTTCCCTTGAAGGCTCCGTGGCGCTCTCCTGTTCGACCTGCGGCTTAC	4628
Db	10542	CCAGGCTTTCCCTTGAAGGCTCCGTGGCGCTCTCCTGTTCGACCTGCGGCTTAC	10601
OY	4629	CGATACCTGTCCGCTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACGCTG	4688
Db	10602	CGATACCTGTCCGCTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACGCTG	10661
OY	4689	TAGGTATCTCAAGTTGGGTGATGTCGTCCTCCAAAGTGGGCTGTGTGCAGAAACCCC	4748
Db	10662	TAGGTATCTCAAGTTGGGTGATGTCGTCCTCCAAAGTGGGCTGTGTGCAGAAACCCC	10721
OY	4749	CGTTACGCCGACCGCTCGGCTTATCCGGTAACTATGCTTGTGATCCAAACCGCGTAG	4808
Db	10722	CGTTACGCCGACCGCTCGGCTTATCCGGTAACTATGCTTGTGATCCAAACCGCGTAG	10781
OY	4809	ACACGACTTATCGCACTGGCAGACGCACTGTGAACAGATTAAGCAGAGGAGTATGT	4868
Db	10782	ACACGACTTATCGCACTGGCAGACGCACTGTGAACAGATTAAGCAGAGGAGTATGT	10841
OY	4869	AGCGGCTCTACAGAGTTCTTGAAGTGTGGCTTAACTAAGGCTACATAGAGGACGT	4928
Db	10842	AGCGGCTCTACAGAGTTCTTGAAGTGTGGCTTAACTAAGGCTACATAGAGGACGT	10901
OY	4929	ATTGGTATCTGCGCTGTGGAAGCCAGTTACCTTGGAAAAAGAGTTGGTAGCTCTTG	4988
Db	10902	ATTGGTATCTGCGCTGTGGAAGCCAGTTACCTTGGAAAAAGAGTTGGTAGCTCTTG	10961
OY	4989	ATCCGGCAAAACACACCGCTGTAGCGGTGTGTTTTTTTGTGCAAGCAGCAGATTAAC	5048

```
|||||
Db 10962 ATCCGGAACAAACCCCGCTGAGCGGGTGTGTTTGTGTTGCAAGCAGAGATTAC 11021
OY GCGGAGAAAAAAGATCTCAAGAAATCCTTTGATCTTTCTACGGGGGTGACGGCA 5108
Db 11022 GCGGAGAAAAAAGATCTCAAGAAATCCTTTGATCTTTCTACGGGGGTGACGGCA 11081
OY GTGGAACGAAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAAAGATCTTAC 5168
Db 11082 GTGGAACGAAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAAAGATCTTAC 11141
OY CTGATCCCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATATGAGTAAAC 5228
Db 11142 CTGATCCCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATATGAGTAAAC 11201
OY TTGGTCGACGTTACCAATGCTTAATCAGTACGAGGACCTTACGAGATCTCTATT 5288
Db 11202 TTGGTCGACGTTACCAATGCTTAATCAGTACGAGGACCTTACGAGATCTCTATT 11261
OY TCGTTCAATCCATAGTTGCTGCTACCCCGTGTAGTAACTACGATACGGAGGCTT 5348
Db 11262 TCGTTCAATCCATAGTTGCTGCTACCCCGTGTAGTAACTACGATACGGAGGCTT 11321
OY ACAATCTGGCCCCAGTCTGCAATGATACCGGACCCACGCTCACCGGCTCCAGATT 5408
Db 11322 ACAATCTGGCCCCAGTCTGCAATGATACCGGACCCACGCTCACCGGCTCCAGATT 11381
OY ATCAGCATTAACACGACGACGAGGAGGCGGACGAGAGTGGTCTCTGCACTTATC 5468
Db 11382 ATCAGCATTAACACGACGACGAGGAGGCGGACGAGAGTGGTCTCTGCACTTATC 11441
OY CGCCTCCATCCAGTCTATTATTTGTTGCCGGAGCTAGATAGTATGTTGCCAGTTAA 5528
Db 11442 CGCCTCCATCCAGTCTATTATTTGTTGCCGGAGCTAGATAGTATGTTGCCAGTTAA 11501
OY TAGTTTGGCAGAGTTGTTGCCATTGCTACAGGATGTGTGTCAGCGTCTGTTGG 5588
Db 11502 TAGTTTGGCAGAGTTGTTGCCATTGCTACAGGATGTGTGTCAGCGTCTGTTGG 11561
OY TATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGAGTTCATGATCCCATGTT 5648
Db 11562 TATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGAGTTCATGATCCCATGTT 11621
OY GTGCAAAAAAGCGGTAGCTCTCGGCTCCGATCGATGTTGTCAGAAATGTTGCCGC 5708
Db 11622 GTGCAAAAAAGCGGTAGCTCTCGGCTCCGATCGATGTTGTCAGAAATGTTGCCGC 11681
OY AGTGTATACATCATGTGTTATGACGACATGATTAATCTTACTGATGCAATCCGT 5768
Db 11682 AGTGTATACATCATGTGTTATGACGACATGATTAATCTTACTGATGCAATCCGT 11741
OY AAGATGCTTTCTGAGCTGTGATGTAACCAAGCATTTGAGAAATGTTGATGCG 5828
Db 11742 AAGATGCTTTCTGAGCTGTGATGTAACCAAGCATTTGAGAAATGTTGATGCG 11801
OY GCGACGAGTTGCTTGGCCCGGCTCAATACGGGATTAATCCGGCCACATACGAGAAC 5888
Db 11802 GCGACGAGTTGCTTGGCCCGGCTCAATACGGGATTAATCCGGCCACATACGAGAAC 11861
OY TTTTAAAGTGTCTCATTTGAAAAAGCTTTCTTGGGGGGAATACTTCAAGATCTTACC 5948
Db 11862 TTTTAAAGTGTCTCATTTGAAAAAGCTTTCTTGGGGGGAATACTTCAAGATCTTACC 11921
OY GCTGTGAGATCCAGTTGATGTAACCACTGTCACACCCAAAGCATTTAGATCTTT 6008
Db 11922 GCTGTGAGATCCAGTTGATGTAACCACTGTCACACCCAAAGCATTTAGATCTTT 11981
OY TACTTTACAGACGCTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGG 6068
Db 11982 TACTTTACAGACGCTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGG 12041
OY AATTAAGGCGACACGGAATGTTAATATCTATCTCTCTTTTCAATATTTATGAG 6128
|||||
```

```
Db 12042 AATTAAGGCGACACGGAATGTTGAATACATCTCTCTTTTCAATATTTGAG 12101
OY 6129 CATTTATCAGGCTATTGTCTCATAGACGAGTATACATTTGAATGATTTAGAAAAATA 6188
Db 12102 CATTTATCAGGCTATTGTCTCATAGACGAGTATACATTTGAATGATTTAGAAAAATA 12161
OY 6189 ACAATATAGGGTTCCGGCACATTTCCCGGAAGATGCCAC 6229
Db 12162 ACAATATAGGGTTCCGGCACATTTCCCGGAAGATGCCAC 12202

RESULT 2
US-09-943-286-3
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02, JP
; CURRENT APPLICATION NUMBER: US/09/943,286
; NUMBER OF SEQ. ID NOS: 9
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: Plasmid.
US-09-943-286-3

Query Match 41.7%; Score 2597.8; DB 10; Length 8933;
Best Local Similarity 74.3%; Pred. No. 0;
Matches 1997; Conservative 642; Mismatches 42; Indels 6; Gaps 1;

OY 1291 CCATGAGATGTAAGAGATATCAGCACTTGAGATGGGGTGGAAATGGGGACCA 1350
Db 5579 CAATGAGATGTAAGAGATATCAGCACTTGAGATGGGGTGGAAATGGGGACCA 5638
OY 1351 TGCTCTGGGATTTGATGATCTGTAGTGTACAGAAAAATTTGGGTCACCGTATT 1410
Db 5639 UGCUCUUGGAGUUGUUGAUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 5698
OY 1411 ATGGGTACCTGTGTGGAAGAACCAACCACTGATTTTGGCATCGATGCTTAAG 1470
Db 5699 AUGGGUACUGUUGGAAAGAACCAACCACTGATTTTGGCATCGATGCTTAAG 5758
OY 1471 CATATGATACAGAGTACATTAATGTTGGGCACACATGCTGTACCCACAGACCCA 1530
Db 5759 CAUUGAUGACAGAGUACAUUUGUUGGACACACAUUGUUGUUGUUGUUGUUGUUGUUG 5818
OY 1531 ACCCACAAGAGTATTTGTTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA 1590
Db 5819 ACCCACAAGAGTATTTGTTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA 5878
OY 1591 TGTGTAGAACAGATCAGAGATTAATAGTTATGATGATGATGATGATGATGATGAT 1650
Db 5879 UGGUAGAACAGUACAGAGATTAATAGTTATGATGATGATGATGATGATGATGATGAT 5938
OY 1651 TAAATATTAACCCCTCTGTGTGTTTAAAGTCACTGATTTGTAAGATGATGATGATGAT 1710
Db 5939 UAAAUUUAACCCCTCTGTGTGTTTAAAGTCACTGATTTGTAAGATGATGATGATGAT 5998
OY 1711 CCAATAGTAGAGCGGAGATGATTAATGAGAAAAAGAGATTAATAAATGCTCTTTCA 1770
Db 5999 CCAUUAUUAACCCCTCTGTGTGTTTAAAGTCACTGATTTGTAAGATGATGATGATGAT 6058
OY 1771 ATATCAGCACAAGATTAAGATTAAGTGCAGAAAAAGATATGCTTTTATTAATCTG 1830
Db 6059 AUUUCAGCACAAGATTAAGATTAAGTGCAGAAAAAGATATGCTTTTATTAATCTG 6118
|||||
```

QY	2905	AGGCGACAGCAATTTATTTGTCTGATATATAGTGCAGCAGCAGCAAGCAAAATTTGCTGAGGGCTATATG	2964
Db	7199	AGGCGACAGCAAAUUAUUUCUGUGAUAUGUUCACAGCAGCAAGAAUUAUUUGUGUGGAGGCGUAUUG	7258
QY	2965	AGGGCGAAACAGCAATCTTTGTCACTACACAGTCTGGGGGCGTCAAAACAGCTCCAGGCGAAGAA	3024
Db	7259	AGGCGCAACAGCAUUCUUCUGCAACUCCACAGUUCUGGGGCGAUCACAGCAGCUCCAGGCAAGAA	7318
QY	3025	TCTGTGCTGTGGAAAGATACCTTAAGAGATCAACAGCTCCTGGGGATTTGGGGTTGCTGTG	3084
Db	7319	UCCUGGCGUUGGAAAGAGUACCUAAAGAUCAACAGCAGCUCCUGGGGAAUUUGGGGUGUCUCUG	7378
QY	3085	GAAACACATTTTGCACACACTGCTGTGCTCTGGAATTCGTAGTTGGGATTAATAATCTGCG	3144
Db	7379	GAAAAUCCAUUUCGACCACUCGUCUGCCUUGGAUUCUAGUGGAGUAUUAUUAUUCUCUGG	7438
QY	3145	AACAGATTTGGAAATTAACATGACCTGGATGGAGTGGAGGACAGAGAATAATTACATTTACACAA	3204
Db	7439	AACGAAUUAUGGAUUAACAUAGACCUGGAUGGAGGAGACAGAGAAAUUAACAAUUAUCACAA	7498
QY	3205	GCTTAATACACTCTCTTAATTTGAAGAATTCGCAAAACAGCAGCAAGAAAGAAATGACAAAGAT	3264
Db	7499	GCUUAACACACUCCUUAUUGAAGAAUCCGAAAAACGCAACAAAAAGAAUAGAACAAAGAU	7558
QY	3265	TATTGGAAATTAAGATTAATGGCGAAGTTTGGGAATTTGGGATTAACATTAACAAATGGCGT	3324
Db	7559	UAUUGGAUUAUGAUAUUAUUGGGCAAGUUGUGGAUUGGUAUUAACAUUAACAAUUAUUGGCGU	7618
QY	3325	GGTATATAAAATTAATTCATATATATATAGTAGAGAGGCTTGTAGTTTAAGAAATAGTTTGTG	3384
Db	7619	GGUUAUAAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	7678
QY	3385	CTGTACTTTCTATATATGTAATAGAGTTTAGCGAGGATATTCCACATTATCTGTTTACAGACC	3444
Db	7679	CUGAACUUCUUGAUGUAUAGUUAUAGUUAAGCAGGGAUUAUCCACCAUUAUUCGUUCAGACCC	7738
QY	3445	ACCTCCCAATCCCGAGGGGAGCCCGACAGGCGCCGAAGGAATTAAGAAGAAAGGTTGAGAGA	3504
Db	7739	ACCTCCCAUUCGAGGGGAGCCCGACAGGCGCCGAAGGAUUAAGAAGAAAGGUGGAGAGA	7798
QY	3505	GAGACAGAGACAGATCCATTTCGATTAGTGAAGAGATCTTTAGCACTTATCTGGAGCATC	3564
Db	7799	GAGACAGAGACAGAUCCAUUCGUUAUUGUUAACGAGACUUAACACUUAUUCUGGGAGCAUC	7858
QY	3565	TGCGGAGCCGTGTGCTTTACGTTACCAACCGCTTTGAGAGACTTACTCTTGATTTGAACGA	3624
Db	7859	UGCGGACCUGUGGCUUCUACGUAUCCACCGCUCUGAGAACUUAUUAUUAUUAUUAUUAUUAU	7918
QY	3625	GGATTGTGGAATCTCTGGAGCGAGGGGGGGGGAACCCCTCAAAATTAATTTGGTGGATATCC	3684
Db	7919	GGAUUUGUGGAUUCUUCUGAGCGCAGAGGGGUGGGAACCCUCAAUAUUAUUGUGGAUUCUC	7978
QY	3685	TACAGATTGTGAGTCAAGAACTTAAGAAATATGTCTTTAATCTTCAATTCGCCACAGCCA	3744
Db	7979	UACAGUAUUAUGAUAUUAAGACAGCAUAAAGAUUAGUUCUUAUUGCUCUAUUGGCCACAGCUA	8038
QY	3745	TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAATATTACAAGCAGTTTATAAGCTA	3804
Db	8039	UAGCAGUAGCUGGAGGGGACAGAAUGGGUUAUAAUAUUAUUAUUAUUAUUAUUAUUAUUAU	8098
QY	3805	TTGCGCCATACCTAGAGAAATTAAGACAGGGCGTTTGAAGAGATTTTGGCTATTAAGTGGGT	3864
Db	8099	UUUGCCCAUUAUUAUUAAGACAGGGGCUUGGAAAGAAAGAUUUUUUUGCUUAUUAAGAUUGGU	8158
QY	3865	GGCAAGTGTGTCAAAAAGTACTGTGATTGATGTGGCTGTGTAAAGGAAAGAAATGAGACGA	3924
Db	8159	GGCAAGUGUGCAAAAGUUAUUGUGUGUGUGUGGCUUGCUUAUUAAGGAAAGAAUUAUGAGACA	8218
QY	3925	GCTGAGCCACAGCAGATGGGGTGGAGCAGATATCTCGAATCTTGA	3971
Db	8219	GCUGAGCCACAGCAGUUGGGUGGAGAGCAUUCUCGAGACCUAGA	8265





QY	2125	ATGCTAAAAACATAATAGTACAGCTGAAACACATCTGTAGAAATTAATTGTACAAGACCCA	2184
Db	6419	AUGCGAAAAACCAUAAUADUACGACGACCAACUUCUGUAGAAAUUAAUUGUACAAGACCCA	6478
QY	2185	ACAACAATACAAAGAAAGTATCCCTATCCAGAGGGGACCGAGGAGACATCTTGTTACAA	2244
Db	6479	ACACACAAUACAGAAAAAGUADUCCGUAUCCAGAGAGACACAGGGAGACAUUUGUACAA	6538
QY	2245	TAGGAAAAATAGGAATATGAGACAAAGCACATTGTACATTAGTAGACAAATAGCAATG	2304
Db	6539	UAGGAAAAAUUAGGAUAUUGAGACAAAGACAUCUUGUAACAUUAGUAGACAAAUUGAUAU	6598
QY	2305	CCACCTTTAAACAGATAGCTAGCTGCAATTAATTAAGAGAACATTTTGAAATTAATTAACAATA	2364
Db	6599	ACACUUDUAAACAGUADUACAGUACAAUUAAGAAACAUAUUGGAAUAUUAUAAACAUAUA	6658
QY	2365	TCTTTAAGCAATCCTCAGAGAGGGACCCGAGAAATTTGTAAACGACAGTTTAAATGTGGAG	2424
Db	6659	UCUUUAAACAGACCTCCAGAGAGGGACCCGAGAAUUGUUAACGACAGUUUUUAUUGUGGAG	6718
QY	2425	GGGAAATTTTCTACGTATATCAACACACATGTTTAATAGTAACTTGCTTAATAGTACTT	2484
Db	6719	GGGAUUUUUUUCCACGUAUUDCAACACACUGUUUAUAGACUUGGUAUUAUAGUACUU	6778
QY	2485	GGAGTACTGAAGGGGCAATTAACATCTGAAGAGAGTGCACATCAACATCTCCATGACGAA	2544
Db	6779	GGAGUACUAAAGGGUCCAUAUAAACUGAAGGAAGUAGACAAUACCCUCCAUAGACAA	6838
QY	2545	TAAACAATTTATTAACATGTGGCAGGAAGTAGGAAAAAGCAATGTATGCCCCCTCCATCA	2604
Db	6839	UAAACCAAAUUUAUAAACUAGUGGGCAGGAAGGAGAAAGACAUUGUADUCCUCCUCCANCA	6898
QY	2605	GTCGACAAATTAGATGTTCATCAATTAATTTACTGGCGCTCTAATTAACAAGAGATGGTGA	2664
Db	6899	GUGGACCAAUUAGAUUGUUCACUCCAAUUAUUAACAGGCGUCUUAUUAACAAGAUUGUGUA	6958
QY	2665	ATAACACAATGGGCTCCGAGATCTTTCAGACCTGGAGAGGCGATATGAGGACAATTGGA	2724
Db	6959	AUAGCAACAAUAGAGUCCGAGAUUCUACAACCUGGAGGAGAGAUUAUGAGGACAAUUGGA	7018
QY	2725	GAAAGCATTTATTAATATATAAGTAGTAAATTAATGAACCTTTGGAGTACCACCCACCA	2784
Db	7019	GAAAGGAUUUAUUAUAAUAUAAUAGUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA	7078
QY	2785	AGGCAAGAGAGAGAGTGGTGACAGAGAGAAAAAGACAGTGGGATATAGAGACTTTGTTCC	2844
Db	7079	AGGCAAAAGAGAGAGUGUGGACAGAGAGAAAAAGAGCGUGGGAAUUAUGAGACUUUUGUCC	7138
QY	2845	TTGGGTTCTTGGGAGCAGCAGAGAACATATGGGCTGCAGCTCAATACGCTGACGATAC	2904
Db	7139	UUGGGUUCUUGGGAGCAGCAGGAAGCACUUAUAGGCGCAGCGUCCAUGACGUGAGCGUAC	7198
QY	2905	AGGCGAGCAATTAATTTGTCTATATAGTACAGCAGACAGAAACAATTTGCTGAGGCGTATTG	2964
Db	7199	AGGCCAGACAAUUAUUGUUGUGUGUUAUGUAGCAGCAGCAAAUUAUUGUGUAGGGGCUUUG	7258
QY	2965	AGGGGCAACAGCATCTGTTGCACTCACAGTCTGGGGCATCAAAACAGCTCCAGGCAAGAA	3024
Db	7259	AGGGGCAACAGCATCUUGUUGCAACUACAGTCCUUGGGGCAUUAACAGCTCCAGGCAAGAA	7318
QY	3025	TCTGTGCGTGTGAAAGATACCTTAAAGGATCAACAGCTCTCGGGGATTTGGGGTGTGCTG	3084
Db	7319	UCCUGGCGUGGAAAGUACCUAAAGAUCAACAGCTCCUGGGGAUUUUGGGGUGUGUCUG	7378
QY	3085	GAAGAACTCATTTGACACACTCTCTGCTTGGAATGCTAGTTGGAGTAAATTAATCTCTGG	3144
Db	7379	GAAGAACTCATTTGACACACTCTCTGCTTGGAATGCTAGTTGGAGTAAATTAATCTCTGG	7438
QY	3145	AACGAGATTTGGAATTAACATGACCTGAGATGAGTGCGGACAGAGAAATTTAACAATTAACAA	3204
Db	7439	AACGAGAUUUGGAAUAAACAUAGCUCUGGUAUGGAGUGGACAGAGAAAUUUAACAUAUACAA	7498

QY		3205	GCTTAATACACTCCTTATTGAGATATCGCAAAAACGACAGAAGAAAGATGAACAAGAT	3264
Dd		7499	GCUTAAUACACUCUCUUUAUUGAAGAACCAGAAAAACGACAAAGAAAGAUUGAACAAAGAU	7558
QY		3265	TATTGGAATTTGATAAATGGGCAAAGTTTGAGAAATTTGTAACTATACAAATTTGGCTGT	3324
Dd		7559	UAUUGGAUUDAGAUAAAUUGGGCAAGUUUGGGAUUUGGUUUAACATACAAAUUGGCCU	7618
QY		3325	GGTATATAAAAATTTTCATATATGATAGTAGAGAGGCTTTGAGGTTTTAAGATATGTTTTG	3384
Dd		7619	GGUUAUAAAAUUUVUADCUAAUUGAUAUAGAGAGCCUGUGAGUUUUAAGAUAUUGUUUG	7678
QY		3385	CTGTACTTCTCATATGTCAATATAGAGTTATAGGAGGATATTCACCATATGCTTTGACACC	3444
Dd		7679	CUGAACUUUCUGUUGUAUUGAGAGUUGGCAAGGAGAUUUDACCUAUUAUGUUUCAGACCC	7738
QY		3445	ACCTCCCAATCCCGAGGAGACCCGACAGGCCCAAGAGAAATAGAGAAGAGGTGAGAGA	3504
Dd		7739	ACCTCCCAAUUCCGAGGAGACCCGACAGGCCCGAGAAUAGAGAAGAGGUGGAGAGA	7798
QY		3505	GAGACAGAGACAGATCCATTTCGATTAGTGAACGGATCCTTAGCACTTATCTGGAGCATC	3564
Dd		7799	GAGACAGAGAGAGAUCCAUUCCGAUUAUGUAAAGCAUCCUAGCACUUAUCUGGGAGAAC	7858
QY		3565	TGCGAGACCTGTGCTCTTACAGTACACACCGCTTGAGAGACCTTACTGATTTGTAACGA	3624
Dd		7859	UGCAGACCCUUGCCUUCUUCAGCCACCCGCUUGAGAGACUUAUCUUGAUUGUAGCA	7918
QY		3625	GGATTGTGAACTTCTGGAGCGCAGGGGTGGGAGCCCTCAAAATATGTGTGAATCTCC	3684
Dd		7919	GGAUUGUGGAAACUUCUGGACGACGAGGGGUGGGAAACCCUCAAAUUAUUGUGUAAUUC	7978
QY		3685	TACAGTATTTGGAGTCAGAGACTAAAGAAATAGTCTCTTAACTTGCTCATATGCCACAGCA	3744
Dd		7979	UACAGUUAUUGGAGGACAGGACUUAAGAUUGUCUUCUUGACUUGCUCUAAUUGCCACAGCA	8038
QY		3745	TAGAGATAGCTGAAGGAGACAGATAGAGGTTTAAAGTAAATTAACAAGCAGCTTATAGAGCTA	3804
Dd		8039	UAGCAGUAGCUGAGAGGACAGAGAUAGGUGUUAUAAGAUAGUACAGAAGAGCUUAUAGAGCUA	8098
QY		3805	TTGGCCACATPACTAGAGAATTAAGACAGGGCTTTGGAAGAAAGATTTTTGCTATAGATGGT	3864
Dd		8099	TUCCCCCAUACUACUAGAAGAAUUAAGACAGGCCUUGGAAAGAUUUUGCUAUAAGAVGGU	8158
QY		3865	GGCAAGGAGUCAAAGAACTAGTGTGATTTGATGGCCGTGTAAGGAAGAAAGATAGAGCA	3924
Dd		8159	GGCAAGGAGUCAAAGAAUAGUGUGUGUGGAGUUGCCUGCGUAAAGAAAGAUUGAGACGA	8218
QY		3925	GCTGAGCCAGCAGCAGATGGGGTGGAGGACAGTATCTCGAGATCTAGA 3971	
Dd		8219	GCUGAGCCAGAGAGCAUUGGGGUGGAGCGACGACUUCUGAGAGACUACA 8265	
 RESULT 5 US-09-737-190A-1 : Sequence 1, Application US/09737190A : Patent No. US20020102545A1 GENERAL INFORMATION: APPLICANT: Shiduya, Tetsuo TITLE OF INVENTION: A Method for Changing a Target Array, a Method for Analyzing TITLE OF INVENTION: A Structure, and an Apparatus, TITLE OF INVENTION: Transmission Medium Therefore FILE REFERENCE: JP919990270US1 (14043) CURRENT APPLICATION NUMBER: US/09/737,190A CURRENT FILING DATE: 2000-12-14 SOFTWARE: PatentIn Ver. 2.1 SEQ ID NO 1 LENGTH: 9719 TYPE: DNA ORGANISM: Human Immunodeficiency virus type 1 US-09-737-190A-1				



Query Match		41.6%;	Score 2593;	DB 10;	Length 9719;
Best Local Similarity		98.1%;	Pred. No. 0;		
Matches 2636;		Conservative	0;	Mismatches	45;
				Indels	6;
				Gaps	1;
OY	1291	CCATGAGAGTGAAGAGAGATATCAGCAGCTGTGAGATGGGGGTGGAATGGGGACCA	1350		
Db	6223	CAATGAGAGTGAAGAGAAATATCAGCAGCTGTGAGATGGGGGTGAGATGGGGACCA	6282		
OY	1351	TGCTCCCTGGGATNTGATGATCTGTAGTGTACAGAAAATTTGGGGTCCACCGCTATT	1410		
Db	6283	TGCTCCCTGGGATNTGATGATCTGTAGTGTACAGAAAATTTGGGGTCCACAGCTATT	6342		
OY	1411	ATGGGCTACTGTGTGAAGAGCAACACCACCTCTATTGTTGTCATCAATGCTAAAG	1470		
Db	6343	ATGGGCTACTGTGTGAAGAGCAACACCACCTCTATTGTTGTCATCAATGCTAAAG	6402		
OY	1471	CATATGATACAGAGGTACATTAATGTTTGGCCACATGCTGTACCCACAGACCCA	1530		
Db	6403	CATATGATACAGAGGTACATTAATGTTTGGCCACATGCTGTACCCACAGACCCA	6462		
OY	1531	ACCCAGAGAGTATGATATGTAATGTGACAGAAAATTTTAACTGTGGAATAATGACA	1590		
Db	6463	ACCCAGAGAGTATGATATGTAATGTGACAGAAAATTTTAACTGTGGAATAATGACA	6522		
OY	1591	TGCTAGACAGATGATGAGATATATCAGTTTATGGGATCAAGCCCTAAAGCCATGTG	1650		
Db	6523	TGCTAGACAGATGATGAGATATATCAGTTTATGGGATCAAGCCCTAAAGCCATGTG	6582		
OY	1651	TAAATTTAACCCACTGTGTGTAGTTTAAAGTGACATGATTTGAAGAATGATTAATA	1710		
Db	6583	TAAATTTAACCCACTGTGTGTAGTTTAAAGTGACATGATTTGAAGAATGATTAATA	6642		
OY	1711	CCATAGATAGTGGGGAGATGATATGAGAAAAGAGATATAAAACCTGCTTTTCA	1770		
Db	6643	CCATAGATAGTGGGGAGATGATATGAGAAAAGAGATATAAAACCTGCTTTTCA	6702		
OY	1771	ATATCAGCACAAGCATAAGAGATAGTGTCAGAAAAGATATGATTTTATAAATCTG	1830		
Db	6703	ATATCAGCACAAGCATAAGAGATAGTGTCAGAAAAGATATGATTTTATAAATCTG	6762		
OY	1831	ATATAGTACCAATAGATA-----TACCAGCTATAGTTGATTAAGTGAACACCTCAG	1884		
Db	6763	ATATAGTACCAATAGATATATATCTACACCTATTAAGTGAACAGTGAACACCTCAG	6822		
OY	1885	TCATTACACAGCCGTGCCAAGGATATCTTTGAGCCCAATTCCTAATATATTTGGCC	1944		
Db	6823	TCATTACACAGCCGTGCCAAGGATATCTTTGAGCCCAATTCCTAATATATTTGGCC	6882		
OY	1945	CGGCTGTTTTGCGATTTCTAAATGTATTAATTAAGACGTTCAATGGAACAGACCATGTA	2004		
Db	6883	CGGCTGTTTTGCGATTTCTAAATGTATTAATTAAGACGTTCAATGGAACAGACCATGTA	6942		
OY	2005	CAAAATGTCAGCAGTACATATGTACACATGGAATTCAGGCGAGTATCAACTCACTGC	2064		
Db	6943	CAAAATGTCAGCAGTACATATGTACACATGGAATTCAGGCGAGTATCAACTCACTGC	7002		
OY	2065	TGTTAAATGGAGCTAGCAGAAAGAGTGTAGTATAGATCTGCCAATTTCCACAGCA	2124		
Db	7003	TGTTAAATGGAGCTAGCAGAAAGAGTGTAGTATAGATCTGCCAATTTCCACAGCA	7062		
OY	2125	ATGCTAAACCATATATAGTACAGTGAACATCTGTGAATTAATTTGTACAAGACCA	2184		
Db	7063	ATGCTAAACCATATATATAGTGAACATCTGTGAATTAATTTGTACAAGACCA	7122		
OY	2185	ACAACATACAGAAAAGATTCGATTCAGAGGGGACACAGGGAGACATTTGTTACAA	2244		
Db	7123	ACAACATACAGAAAAGATTCGATTCAGAGGGGACACAGGGAGACATTTGTTACAA	7182		
OY	2245	TAGGAAAAATAGGAATATGAGACAGCAATTTGTAACATATAGTGAACAAATGGAATG	2304		
Db	7183	TAGGAAAAATAGGAATATGAGACAGCAATTTGTAACATATAGTGAACAAATGGAATG	7242		
OY	2305	CCACTTTAAAAAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAAATTAATTAACAATA	2364		

Db	7243	ACACTTTAAACAGATAGCTAGCAATTAAGAGAACAAATTTGGAATATATAAACAAATA	7302		
OY	2365	TCTTTAAGCAATCTCAGAGAGGACCCAGAAATTTGAACGACATTTTAATTTGGAG	2424		
Db	7303	TCTTTAAGCAATCTCAGAGAGGACCCAGAAATTTGAACGACATTTTAATTTGGAG	7362		
OY	2425	GGGAATTTTCTAGCTAATTCACACAACTGTTTAATAGTACTGTTTAATATACTACT	2484		
Db	7363	GGGAATTTTCTAGCTAATTCACACAACTGTTTAATAGTACTGTTTAATATACTACT	7422		
OY	2485	GGAGTACTGAAGGGTCAATATACTGAAGAGAGACACATCACTCCATCCAGAA	2544		
Db	7423	GGAGTACTGAAGGGTCAATATACTGAAGAGAGACACATCACTCCATCCAGAA	7482		
OY	2545	TAAACAAATTTAATAACATGTGGCAGAGATAGAAAACAAATGTATCCCTCCATCA	2604		
Db	7483	TAAACAAATTTAATAACATGTGGCAGAGATAGAAAACAAATGTATCCCTCCATCA	7542		
OY	2605	GTGGCAATTAAGATGTCATCAATATTTACTGGGCTGCTATTACAGAGATGGTGTA	2664		
Db	7543	GTGGCAATTAAGATGTCATCAATATTTACTGGGCTGCTATTACAGAGATGGTGTA	7602		
OY	2665	ATAACAAATGGGTCCGAGATCTTCAGACCTGGAGAGCGATATGAGGCAATTTGGA	2724		
Db	7603	ATAACAAATGGGTCCGAGATCTTCAGACCTGGAGAGCGATATGAGGCAATTTGGA	7662		
OY	2725	GAAGTAATTTATTAATATTAAGATAGTAAAAATTTGAACATTTAGAGTACCCACCA	2784		
Db	7663	GAAGTAATTTATTAATATTAAGATAGTAAAAATTTGAACATTTAGAGTACCCACCA	7722		
OY	2785	AGGCAAAAGAAAGATGGTGCAGAGAGAAAAAGCAGTGGGAATAGAGCTTTGTCC	2844		
Db	7723	AGGCAAAAGAAAGATGGTGCAGAGAGAAAAAGCAGTGGGAATAGAGCTTTGTCC	7782		
OY	2845	TTGGGTTTGGGAGCAGCAGAGAACACTATGCGCTGCACGTCATGACGCTGACGTTAC	2904		
Db	7783	TTGGGTTTGGGAGCAGCAGAGAACACTATGCGCGCAGCCCATGACGCTGACGTTAC	7842		
OY	2905	AGGCCAGCAATTTATTTCTGATATATGTGACAGCAGACAGAACATTTGCTAGGCTATTG	2964		
Db	7843	AGGCCAGCAATTTATTTCTGATATATGTGACAGCAGACAGAACATTTGCTAGGCTATTG	7902		
OY	2965	AGGGCAACAGACATGTTGCAACACAGTCTGGGGGATCAACAGCTCCAGGCAAGAA	3024		
Db	7903	AGGGCAACAGACATGTTGCAACACAGTCTGGGGGATCAACAGCTCCAGGCAAGAA	7962		
OY	3025	TCTGGGCTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGAATTTGGGTTGCTCTG	3084		
Db	7963	TCTGGGCTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGAATTTGGGTTGCTCTG	8022		
OY	3085	GAATACTATTTGCACACCTGCTGCTTGGGAATGTAGTTGGAATTAATCTCTGG	3144		
Db	8023	GAATACTATTTGCACACCTGCTGCTTGGGAATGTAGTTGGAATTAATCTCTGG	8082		
OY	3145	AACAGATTTGGAAATTAACATGACCTGAGTGGAGTGGGAAGAAATTAACAATTACAA	3204		
Db	8083	AACAGATTTGGAAATTAACATGACCTGAGTGGAGTGGGAAGAAATTAACAATTACAA	8142		
OY	3205	GCTTAATACACTCTTAATTAAGAAATGCAAAAACACAGAAAGAAATGAACAAAGAT	3264		
Db	8143	GCTTAATACACTCTTAATTAAGAAATGCAAAAACACAGAAAGAAATGAACAAAGAT	8202		
OY	3265	TATTGGAATTAAGATTAATGGCAAGTTTGGGAATTTGTTTAACATTAACAAATTTGGCTGT	3324		
Db	8203	TATTGGAATTAAGATTAATGGCAAGTTTGGGAATTTGTTTAACATTAACAAATTTGGCTGT	8262		
OY	3325	GGTATATTAATTAATTAATTAATAGTAGAGAGGCTTGCTAGTTAAGATTAATTTTG	3384		
Db	8263	GGTATATTAATTAATTAATTAATAGTAGAGAGGCTTGCTAGTTAAGATTAATTTTG	8322		
OY	3385	CTGTACTTTCTATAGTGAATAGAGTTAGCAGGGAATATTCACCAATTAATGTTTCAGACC	3444		

Db 8323 CTGTACTTCTTATAGTGAATAGTATAGGACGGGATATTCACCTATTCGTTTCAGACCC 8382  
Qy 3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCCGGAAGAAATAGAAAGAGGTGGAGAGA 3504  
Db 8383 ACCTCCCAATCCCGAGGGGACCCGACAGGCCGGAAGAAATAGAAAGAGGTGGAGAGA 8442  
Qy 3505 GAGACAGAGACAGATTCATTCGATTAAGTGAACGGATCTTACACTTATCTGGGACATC 3564  
Db 8443 GAGACAGAGACAGATTCATTCGATTAAGTGAACGGATCTTACACTTATCTGGGACATC 8502  
Qy 3565 TCGCGAGCCGTGGCTCTTACAGTACACCGCTTGAAGAGATCTCTGATTTGAACGA 3624  
Db 8503 TCGCGAGCCGTGGCTCTTACAGTACACCGCTTGAAGAGATCTCTGATTTGAACGA 8562  
Qy 3625 GGATTTGTGAACCTCTGGGACGACAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCC 3684  
Db 8563 GGATTTGTGAACCTCTGGGACGACAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCC 8622  
Qy 3685 TACAGTATGTGAGTCAGGAACTAAAGAAATAGTGTCTGTTAACTGCTCAATGCCACGCA 3744  
Db 8623 TACAGTATGTGAGTCAGGAACTAAAGAAATAGTGTCTGTTAACTGCTCAATGCCACGCA 8682  
Qy 3745 TACAGTATGTGAGGGGACAGATAGGGTTATAGAAGTATTTACAGCAGCTTATAGGCTA 3804  
Db 8683 TACAGTATGTGAGGGGACAGATAGGGTTATAGAAGTATTTACAGCAGCTTATAGGCTA 8742  
Qy 3805 TTCCGCAACATACCTAGAAAGATAAGACAGGGCTTGAAGAGATTTGCTATATAGATGGGT 3864  
Db 8743 TTCCGCAACATACCTAGAAAGATAAGACAGGGCTTGAAGAGATTTGCTATATAGATGGGT 8802  
Qy 3865 GGCAAGTGTCTAAAAAATAGTGTATGTGATGGCTGCTGTAAGGAAAAAGATAGACGA 3924  
Db 8803 GGCAAGTGTCTAAAAAATAGTGTATGTGATGGCTGCTGTAAGGAAAAAGATAGACGA 8862  
Qy 3925 GCTAGAGCAGCAGATAGGGGTGGAGCAGTATCTGAGATCTAGA 3971  
Db 8863 GCTAGAGCAGCAGATAGGGGTGGAGCAGTATCTGAGATCTAGA 8909

RESULT 6  
US-10-000-511A-1  
: sequence 1, Application US/10000511A  
: Publication No. US20030158131A1  
: GENERAL INFORMATION:  
: APPLICANT: Aldevin1, Anna  
: TITLE OF INVENTION: DNA VECTORS CONTAINING MUTATED HIV PROVRUSES  
: FILE REFERENCE: 101353-165  
: CURRENT APPLICATION NUMBER: US/10/000, 511A  
: CURRENT FILING DATE: 2001-10-23  
: PRIOR APPLICATION NUMBER: 60/242, 589  
: PRIOR FILING DATE: 2000-10-23  
: PRIOR APPLICATION NUMBER: 60/243, 432  
: PRIOR FILING DATE: 2000-11-28  
: NUMBER OF SEQ ID NOS: 37  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 1  
: LENGTH: 9719  
: TYPE: DNA  
: ORGANISM: Human immunodeficiency virus  
US-10-000-511A-1

Query Match 41.68; Score 2593; DB 12; Length 9719;  
Best Local Similarity 98.18; Pred. No. 0;  
Matches 2636; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

Qy 1291 CCATGAGAGTGAAGAGAGATAGACACTTTGTGAAGATGGGGTGGAAATGGGGACCA 1350  
Db 6223 CAATGAGAGTGAAGAGAGATATACGACTTTGTGAAGATGGGGTGGAGATGGGGACCA 6282  
Qy 1351 TGTCTCTGGATATGTATGATCTGTAGTGTACAGAAAAATTTGGGTCAACCTCTATT 1410  
Db 6283 TGTCTCTGGATATGTATGATCTGTAGTGTACAGAAAAATTTGGGTCAACCTCTATT 6342

Qy 1411 ATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTGTCATCAGATCTAAG 1470  
Db 6343 ATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTGTCATCAGATCTAAG 6402  
Qy 1471 CATATGATACAGAGTACATTAATGTTGGGCGACACATGGCTGTGTCCACAGACCCA 1530  
Db 6403 CATATGATACAGAGTACATTAATGTTGGGCGACACATGGCTGTGTCCACAGACCCA 6462  
Qy 1531 ACCCACAAAGATAGTATTTGTAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA 1590  
Db 6463 ACCCACAAAGATAGTATTTGTAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA 6522  
Qy 1591 TGTAGAACAGATCATGAGATTAATCACTTTATGGAATCAAGCCTTAAGCCATGTG 1650  
Db 6523 TGTAGAACAGATCATGAGATTAATCACTTTATGGAATCAAGCCTTAAGCCATGTG 6582  
Qy 1651 TAAATTAACCCCTCTGTGTGTTTAAAGTSCATGTATTTGAAGAAATGATCTATA 1710  
Db 6583 TAAATTAACCCCTCTGTGTGTTTAAAGTSCATGTATTTGAAGAAATGATCTATA 6642  
Qy 1711 CCAATAGTATAGCGGAGAGATGATAATGAGAAAGAGAGATAAAAAAGTCTCTTCA 1770  
Db 6643 CCAATAGTATAGCGGAGAGATGATAATGAGAAAGAGAGATAAAAAAGTCTCTTCA 6702  
Qy 1771 ATATCAGACAGACATTAAGATTAAGTGTGACAGAAAGATATGCTTTTATTAACCTG 1830  
Db 6703 ATATCAGACAGACATTAAGATTAAGTGTGACAGAAAGATATGCTTTTATTAACCTG 6762  
Qy 1831 ATATAGTACCAATATGATAA-----TACAGCTTATGATGATGATGTTGATACACCTGAG 1884  
Db 6763 ATATATATACCAATATGATAATGATATCTACACCTTATGATGATGATGTTGATACACCTGAG 6822  
Qy 1885 TCATTAACAGAGCCCTGTCCAAAGGATCTTTGAGCCAAATTCCTCATATTATTTGCCCC 1944  
Db 6823 TCATTAACAGAGCCCTGTCCAAAGGATCTTTGAGCCAAATTCCTCATATTATTTGCCCC 6882  
Qy 1945 CGGCTGTTTGGCATCTTAATATGTAATTAAGACCTTCAATGGAACGACCATGTA 2004  
Db 6883 CGGCTGTTTGGCATCTTAATATGTAATTAAGACCTTCAATGGAACGACCATGTA 6942  
Qy 2005 CAAATGTACAGACAGTACATGTATGATGGAATGGAATCAGGCCAGTATGATCAATCACTGTC 2064  
Db 6943 CAAATGTACAGACAGTACATGTATGATGGAATGGAATGGAATGGAATGGAATGGAATGTC 7002  
Qy 2065 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAATTAAGATGTCGCAATTTCCACAGACA 2124  
Db 7003 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAATTAAGATGTCGCAATTTCCACAGACA 7062  
Qy 2125 ATGCTAAACCATATATATGACAGTGAACACATCTGTGAATTAATTTGAACAAGACCA 2184  
Db 7063 ATGCTAAACCATATATATGACAGTGAACACATCTGTGAATTAATTTGAACAAGACCA 7122  
Qy 2185 ACAACATPACAGAAAAAGATATCCGTATCCAGAGGAGACCAAGGAGAGACATTTGTTCAA 2244  
Db 7123 ACAACATPACAGAAAAAGATATCCGTATCCAGAGAGACCAAGGAGAGACATTTGTTCAA 7182  
Qy 2245 TAGGAAAAATAGGAAATATGAGACAGACATTTGTAACATTAAGAGACAAATGAGATG 2304  
Db 7183 TAGGAAAAATAGGAAATATGAGACAGACATTTGTAACATTAAGAGACAAATGAGATA 7242  
Qy 2305 CCATTTAAACAGATATGCTAGCAATTAAGAGACAAATTTGGAATTAATTAACAATA 2364  
Db 7243 ACATTTAAACAGATATGCTAGCAATTAAGAGACAAATTTGGAATTAATTAACAATA 7302  
Qy 2365 TCTTTAGCAATCCTCAGAGGGGAGCCAGAAATTTGAGCAGCTTAAATTTGGAG 2424  
Db 7303 TCTTTAGCAATCCTCAGAGGGGAGCCAGAAATTTGAGCAGCTTAAATTTGGAG 7362  
Qy 2425 GGGAAATTTTCTACTGTAATTAACACACAGCTTTTAACTACTTGTGTTAAATAGTACT 2484  
Db 7363 GGGAAATTTTCTACTGTAATTAACACACAGCTTTTAACTACTTGTGTTAAATAGTACT 7422  
Qy 2485 GGAATGTAAGGGTCAATTAACACTGAGGAAGTGAACAAATCACACTCCATGACGAA 2544

```
|||||
Db 7423 GGAGTACTGAAGGGTCAAAATTAACACTGAGAGAGTGAACACATCAACCTCCATGACAA 7482
Oy 2545 TAAACAAATTTATTAACATGTGGCAGAGATAGAAAAACAATGTATGCCCTCCATCA 2604
Db 7483 TAAACAAATTTATTAACATGTGGCAGAGATAGAAAAACAATGTATGCCCTCCATCA 7542
Oy 2605 GTGACAAATTTAGATGTTCAATCAATATTTACTGGGCTGCTTTAAACAGAGATGTGTGA 2664
Db 7543 GTGACAAATTTAGATGTTCAATCAATATTTACTGGGCTGCTTTAAACAGAGATGTGTGA 7602
Oy 2665 ATAAACAATGGGTCCGAGATCTTCAGACCTTGAGAGAGGCGATATGAGGACCAATTGGA 2724
Db 7603 ATAGCAACAATGAGTCCGAGATCTTCAGACCTTGAGAGAGGAGATATGAGGACCAATTGGA 7662
Oy 2725 GAAGTAAATTTATTAATATTAATTAAGTAAATTTGAACCATTTAGAGTACACCCACCA 2784
Db 7663 GAAGTAAATTTATTAATATTAATTAAGTAAATTTGAACCATTTAGAGTACACCCACCA 7722
Oy 2785 AGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGACAGTGGGAATAGACCTTTGTC 2844
Db 7723 AGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGACAGTGGGAATAGACCTTTGTC 7782
Oy 2845 TTGGGTTCTTGGAGAGCAGCAGAGAACTATGGCTGCACGTCAATGACGCTTGACGTAC 2904
Db 7783 TTGGGTTCTTGGAGAGCAGCAGAGAACTATGGCTGCACGTCAATGACGCTTGACGTAC 7842
Oy 2905 AGGCCAGACATTTATGCTGATATATGTCAGACAGCAGACAAATTTGCTGAGGCTATTTG 2964
Db 7843 AGGCCAGACATTTATGCTGATATATGTCAGACAGCAGACAAATTTGCTGAGGCTATTTG 2964
Oy 2965 AGGGCAACAGCATCTGTGCACTACAGTCTGGGCAATTAACAGCTCCAGGCAAGAA 3024
Db 7903 AGGGCAACAGCATCTGTGCACTACAGTCTGGGCAATTAACAGCTCCAGGCAAGAA 3024
Oy 3025 TCCTGGCTGTGGAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGCTGCTG 3084
Db 7963 TCCTGGCTGTGGAAGATACCTAAAGATCAACAGCTCTGGGATTTGGGCTGCTG 8022
Oy 3085 GAAACATCATTTGCAACACCTGCTGCTGGTGAATGCTAGTTGAGATTAATCTCTGG 3144
Db 8023 GAAACATCATTTGCAACACCTGCTGCTGGTGAATGCTAGTTGAGATTAATCTCTGG 8082
Oy 3145 AACGATTTGGAATTAACATGACCTGAGTGGAGTGGGACAGGAATTAACATTAACACAA 3204
Db 8083 AACGATTTGGAATTAACATGACCTGAGTGGAGTGGGACAGGAATTAACATTAACACAA 8142
Oy 3205 GCTTAATACACTCCTTAATTAAGAAATCGCAAAACAGCAGAAAGAAATGACAGAAAT 3264
Db 8143 GCTTAATACACTCCTTAATTAAGAAATCGCAAAACAGCAGAAAGAAATGACAGAAAT 8202
Oy 3265 TATTGGAATTTAGATTAATGGGCAAGTTTGGAAATGTTGTAACATAACAAATGGCTGT 3324
Db 8203 TATTGGAATTTAGATTAATGGGCAAGTTTGGAAATGTTGTAACATAACAAATGGCTGT 8262
Oy 3325 GGTATTAATAAATTTATCATATATGATAGTAGAGGCTGCTAGTTTAAGATTAATTTTG 3384
Db 8263 GGTATTAATAAATTTATCATATATGATAGTAGAGGCTGCTAGTTTAAGATTAATTTTG 8322
Oy 3385 CTGACTTCTTATAGTAAATAGATTAAGCAGGATATTCACATTTATCTTTAGACCC 3444
Db 8323 CTGACTTCTTATAGTAAATAGATTAAGCAGGATATTCACATTTATCTTTAGACCC 8382
Oy 3445 ACCTCCCAATCCGAGGGGACCCGACAGGCGCGAAGAAATAGAAAGAAAGGTGAGAGA 3504
Db 8383 ACCTCCCAATCCGAGGGGACCCGACAGGCGCGAAGAAATAGAAAGAAAGGTGAGAGA 8442
Oy 3505 GAGACAGAGACAGATTCGATTTAGTAAAGGATGCTTATGACCTTATCGGAGAGATC 3564
Db 8443 GAGACAGAGACAGATTCGATTTAGTAAAGGATGCTTATGACCTTATCGGAGAGATC 8502
Oy 3565 TGGGAGGCTGCTGCTTACAGTACCAACCCCTTGAGAGATTAATCTTTGATTTAAGCA 3624
|||||
```

```
Db 8503 TGGGAGGCTGTGCTCTTACAGTACCAACCGCTTGAGAGACTTACTCTTGATTTAAGCA 8562
Oy 3625 GGATTTGGAACCTTCTGGGACGCGAGGGGCTGGGAAGCCCTTAATTTGTTGGAATCTCC 3684
Db 8563 GGATTTGGAACCTTCTGGGACGCGAGGGGCTGGGAAGCCCTTAATTTGTTGGAATCTCC 8622
Oy 3685 TACAGTATTTGGAGTCAGAACTAAAGAAATAGTGTCTTACTTGTCTCAATCCACAGCA 3744
Db 8623 TACAGTATTTGGAGTCAGAACTAAAGAAATAGTGTCTTACTTGTCTCAATCCACAGCA 8682
Oy 3745 TAGCAGTAGCTGAGGGGACACATAGGGCTTATAGAATTTCAAGCACTTATAGAGCTA 3804
Db 8683 TAGCAGTAGCTGAGGGGACACATAGGGCTTATAGAATTTCAAGCACTTATAGAGCTA 8742
Oy 3805 TTGCCACATACCTAGAAAGAAATTAAGACAGGCTTGAAAGGATTTGCTATTAAGTGGGT 3864
Db 8743 TTGCCACATACCTAGAAAGAAATTAAGACAGGCTTGAAAGGATTTGCTATTAAGTGGGT 8802
Oy 3865 GGCAAGTGGTCAAAAAGTAGTGTGATTTGATTTGGCTGCTGTTAAGGAAAGATGAGAGA 3924
Db 8803 GGCAAGTGGTCAAAAAGTAGTGTGATTTGATTTGGCTGCTGTTAAGGAAAGATGAGAGA 8862
Oy 3925 GCTGAGCCAGCAGCAGATGAGGTGGAGAGCACTATCTGAGATCTAGA 3971
Db 8863 GCTGAGCCAGCAGCAGATGAGGTGGAGAGCACTATCTGAGACCTTGA 8909

RESULT 7
US-10-000-511A-2
: Sequence 2, Application US/10000511A
: Publication No. US20030158131A1
: GENERAL INFORMATION:
: APPLICANT: Algovin, Anna
: TITLE OF INVENTION: DNA VECTORS CONTAINING MUTATED HIV PROVIRUSES
: FILE REFERENCE: 101353-165
: CURRENT APPLICATION NUMBER: US/10/000, 511A
: CURRENT FILING DATE: 2001-10-23
: PRIOR APPLICATION NUMBER: 60/242, 589
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/243, 432
: PRIOR FILING DATE: 2000-11-28
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 9719
: TYPE: DNA
: ORGANISM: Human Immunodeficiency virus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(9717)
: US-10-000-511A-2

Query Match 41.6%; Score 2593; DB 12; Length 9719;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 45; Indels 6; Gaps 1;
```

|||||  
Db 6463 ACCACACAGAGATGTAATGTGTAATGTGACAGAAAATTTTAACATGTGGAAAAATACACA 6522  
QY 1591 TGGTAGAACAGATGTCATGAGATATTAATCAGTTTATGAGATCAAAAGCCTAAAGCCATGTG 1650  
Db 6523 TGGTAGAACAGATGTCATGAGATATTAATCAGTTTATGAGATCAAAAGCCTAAAGCCATGTG 6582  
QY 1651 TAAAATTAACCCCATCTGTGTATAGTTTAAAGTGCACCTGATTTGACAAATGATATCTATA 1710  
Db 6583 TAAAATTAACCCCATCTGTGTATAGTTTAAAGTGCACCTGATTTGACAAATGATATCTATA 6642  
QY 1711 CCAATAGTAGAGGGAGATGATTAATGGAGAAAGAGATAAAAACGCTTGTCA 1770  
Db 6643 CCAATAGTAGAGGGAGATGATTAATGGAGAAAGAGATAAAAACGCTTGTCA 6702  
QY 1771 ATATCAGACAGAGCATTAAGATTAAGTGCAGAAAGATATGCATCTTTTATTAACCTG 1830  
Db 6703 ATATCAGACAGAGCATTAAGATTAAGTGCAGAAAGATATGCATCTTTTATTAACCTG 6762  
QY 1831 ATATAGTACCAATATGATTA-----TACAGCTATAGCTTGATATAGTTGTAACACCTCAG 1884  
Db 6763 ATATTAATTAACCAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 6822  
QY 1885 TCATTAACAGAGCCTGTGCAAGGATCCTTTGAGCCAAATTCATCATTAATGTGCCC 1944  
Db 6823 TCATTAACAGAGCCTGTGCAAGGATCCTTTGAGCCAAATTCATCATTAATGTGCCC 6882  
QY 1945 CGGCTGTTTTGCGATTTCTAAAATGTATATTAAGAGCTTCAATGAGAACAGGACATGTA 2004  
Db 6883 CGGCTGTTTTGCGATTTCTAAAATGTATATTAAGAGCTTCAATGAGAACAGGACATGTA 6942  
QY 2005 CAAATGTCAGACAGTACAAATGTAACATGGAATCAGGCCAGTATGATCAATCACTACCTGC 2064  
Db 6943 CAAATGTCAGACAGTACAAATGTAACATGGAATCAGGCCAGTATGATCAATCACTACCTGC 7002  
QY 2065 TGTTAATGGAGCTGTGCAAGGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2124  
Db 7003 TGTTAATGGAGCTGTGCAAGGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 7062  
QY 2125 ATGCTAAAACCATTAATAGTACAGCTGTAACACATCTGTAGAAATTAATTTGTACAAAGACCA 2184  
Db 7063 ATGCTAAAACCATTAATAGTACAGCTGTAACACATCTGTAGAAATTAATTTGTACAAAGACCA 7122  
QY 2185 ACAACATATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTCAA 2244  
Db 7123 ACAACATATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTCAA 7182  
QY 2245 TAGGAAAAATAGGAAAAATAGAGCAACACATTTGTATAGTATAGTATAGTATAGTATAGTATAGTAT 2304  
Db 7183 TAGGAAAAATAGGAAAAATAGAGCAACACATTTGTATAGTATAGTATAGTATAGTATAGTATAGTAT 7242  
QY 2305 CCACTTTAAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAAAATTAATAACAATAA 2364  
Db 7243 ACACTTTAAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAAAATTAATAACAATAA 7302  
QY 2365 TCTTTAAGCAATCTCTCGAGAGGGGACCCAGAAATTTGTAACGACACAGTTTAAATTTGGAG 2424  
Db 7303 TCTTTAAGCAATCTCTCGAGAGGGGACCCAGAAATTTGTAACGACACAGTTTAAATTTGGAG 7362  
QY 2425 GGGAAATTTTCTACTGTAAATTCACACACACTGTTTATAGTACTGTTTAAATTTGGAG 2484  
Db 7363 GGGAAATTTTCTACTGTAAATTCACACACACTGTTTATAGTACTGTTTAAATTTGGAG 7422  
QY 2485 GAGTACTGAGGGTCAAAATTAACACTGAAGAGTGCACAAATCAGCTCCCATCGAGAA 2544  
Db 7423 GAGTACTGAGGGTCAAAATTAACACTGAAGAGTGCACAAATCAGCTCCCATCGAGAA 7482  
QY 2545 TAAAACATTTATTAACATGTGCGAGAGAGTATAGAAAAAGCAATGTATGCCCTCCATCA 2604  
Db 7483 TAAAACATTTATTAACATGTGCGAGAGAGTATAGAAAAAGCAATGTATGCCCTCCATCA 7542  
QY 2605 GTGAGCAAAATTAAGTGTTCATCAAAATTTACTGGGCTGTAATTAACAGAGATGCTGTGTA 2664  
|||||

Db 7543 GTGAGCAAAATTAAGTGTTCATCAAAATTTTACAGGCTGCTATTTAACAGAGATGCTGTGTA 7602  
QY 2665 ATAAACAACAATGGGTCCGAGATCTTCAGACCTGAGAGGAGCGATATAGGAGCAATTTGA 7724  
Db 7603 ATAGCAACAACAATGGGTCCGAGATCTTCAGACCTGAGAGGAGGATATAGGAGCAATTTGA 7662  
QY 2725 GAAATGAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2784  
Db 7663 GAAATGAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7722  
QY 2785 AGGCAAAAGAGAGAGTGTGTCAGAGAGAAAAAGAGAGAGGGAATTTGAGAGCTTTGTTCC 2844  
Db 7723 AGGCAAAAGAGAGAGTGTGTCAGAGAGAAAAAGAGAGAGGGAATTTGAGAGCTTTGTTCC 7782  
QY 2845 TTGGGTTCTTTGGAGCAGAGAGGAGGACCTATGGGCTGCACGTCATATGACCTGACGCTAC 2904  
Db 7783 TTGGGTTCTTTGGAGCAGAGAGGAGGACCTATGGGCTGCACGTCATATGACCTGACGCTAC 7842  
QY 2905 AGGCCAGACAAATTTATTTGCTGATTAATGTCAGAGCAGAGAAACAATTTGCTAGAGGCTATTG 2964  
Db 7843 AGGCCAGACAAATTTATTTGCTGATTAATGTCAGAGCAGAGAAACAATTTGCTAGAGGCTATTG 7902  
QY 2965 AGGGGCAACAGCATCTGTGCAACTGCACAGCTGAGGGGCAATCAACAGCTCCAGGCAAGAA 3024  
Db 7903 AGGGGCAACAGCATCTGTGCAACTGCACAGCTGAGGGGCAATCAACAGCTCCAGGCAAGAA 7962  
QY 3025 TCCGTGCTGTGAGAAAGATTAACCTAAAGGATCAACAGCTCCAGGGGATTTGGGTTGCTCTG 3084  
Db 7963 TCCGTGCTGTGAGAAAGATTAACCTAAAGGATCAACAGCTCCAGGGGATTTGGGTTGCTCTG 8022  
QY 3085 GAAAACCTCATTTGACACCTGCTGTGCTTTGGAATGCTAGTTGAGTAAATTAATCTCTG 3144  
Db 8023 GAAAACCTCATTTGACACCTGCTGTGCTTTGGAATGCTAGTTGAGTAAATTAATCTCTG 8082  
QY 3145 AACAGATTTGGAATTAATGATGACCGGATGAGTGGAGGAGCAGAGAAATTAACAATTAACAA 3204  
Db 8083 AACAGATTTGGAATTAATGATGACCGGATGAGTGGAGGAGCAGAGAAATTAACAATTAACAA 8142  
QY 3205 GCTTAATTAACACTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3264  
Db 8143 GCTTAATTAACACTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8202  
QY 3265 TATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3324  
Db 8203 TATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8382  
QY 3325 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3384  
Db 8383 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8442  
QY 3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCGCCCAAGAAATTAAGAAAGAGGTGGAGAGA 3504  
Db 8443 ACCTCCCAATCCCGAGGGGACCCGACAGGCGCCCAAGAAATTAAGAAAGAGGTGGAGAGA 8442  
QY 3505 GAGACAGAGCAGATCCATTCGATTTAGTGAAGCGATCTTATAGCACTTATGAGGAGATC 3564  
Db 8443 GAGACAGAGCAGATCCATTCGATTTAGTGAAGCGATCTTATAGCACTTATGAGGAGATC 8502  
QY 3565 TGGGAGAGCCTGTGCTCTTCAGAGTACACCGCTTGAGAGAGACTTACTTGTATTTGAAGCA 3624  
Db 8503 TGGGAGAGCCTGTGCTCTTCAGAGTACACCGCTTGAGAGAGACTTACTTGTATTTGAAGCA 8562  
QY 3625 GGAATTTGGAGACTTTGAGAGCAGAGGGGTTGGAGAGCCCTCAAAATTTGTTGGAATTTCTC 3684  
Db 8563 GGAATTTGGAGACTTTGAGAGCAGAGGGGTTGGAGAGCCCTCAAAATTTGTTGGAATTTCTC 8622  
QY 3685 TACAGTATTTGAGATGAGAGCACTAAAGAAATGCTGTTAACTGTCATATGAGGAGGAGCA 3744  
Db 8623 TACAGTATTTGAGATGAGAGCACTAAAGAAATGCTGTTAACTGTCATATGAGGAGGAGCA 8682  
|||||

OY	3745	TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAATTTACAAGACGTTTATAGACCTA	3804
Db	8683	TAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAATTTACAAGACGTTTATAGACCTA	8742
OY	3805	TTGGCCACATACCCATGAGAAATATAGACAGGGCTTGAAAGGATTTGCTTAATAGATGGGT	3864
Db	8743	TTGGCCACATACCCATGAGAAATATAGACAGGGCTTGAAAGGATTTGCTTAATAGATGGGT	8802
OY	3865	GGCAAGTGGTCAAAAAGTAGTGTGATTTGATGGCCTCTGTAAAGGAAAGATGAGACGA	3924
Db	8803	GGCAAGTGGTCAAAAAGTAGTGTGATTTGATGGCCTCTGTAAAGGAAAGATGAGACGA	8862
OY	3925	GCTAGGCACAGACAGATGGGGTGGAGACAGTATCTGAGATCTAGA	3971
Db	8863	GCTAGGCACAGACAGATGGGGTGGAGACAGTATCTGAGATCTAGA	8909
RESULT 8			
US-10-286-332A-14			
: Sequence 14, Application US/10286332A			
: Publication No. US20030138459A1			
: GENERAL INFORMATION:			
: APPLICANT: Wang, Danher			
: TITLE OF INVENTION: METHOD OF VACCINATION THROUGH SEROTYPE ROTATION			
: FILE REFERENCE: 22488-738			
: CURRENT FILING DATE: 2003-03-17			
: PRIOR APPLICATION NUMBER: US 09/585, 599			
: PRIOR FILING DATE: 2000-06-02			
: PRIOR APPLICATION NUMBER: PCT/US01/18238			
: PRIOR FILING DATE: 2001-06-04			
: PRIOR APPLICATION NUMBER: US 10/003, 035			
: PRIOR FILING DATE: 2001-11-01			
: NUMBER OF SEQ ID NOS: 75			
: SOFTWARE: PatentIn version 3.1			
: SEQ ID NO 14			
: LENGTH: 3157			
: TYPE: DNA			
: ORGANISM: Artificial sequence			
: FEATURE:			
: OTHER INFORMATION: Env/Tat/Rev			
US-10-286-332A-14			
Query Match 41.6%: Score 2592.4: DB 12: Length 3157:			
Best Local Similarity 98.2%: Pred. No. 0:			
Matches 2633: Conservative 0: Mismatches 41: Indels 6: Gaps 1:			
OY	1291	CCATGAGAGTGAAGAGAAAGTATCATGCACTTGTGAGATGGGGGTGGAATGGGGCACCA	1350
Db	478	CAATGAGAGTGAAGAGAAATATCATGCACTTGTGAGATGGGGGTGGAATGGGGCACCA	537
OY	1351	TGCTCCCTTGGGATTTGATGATCTGTGATGCTTACAGAAAAATTTGGGTCAAGCTTAATT	1410
Db	538	TGCTCCCTTGGGATTTGATGATCTGTGATGCTTACAGAAAAATTTGGGTCAAGCTTAATT	597
OY	1411	ATGGGGTACCCGTGTGGAAGAACCAACCACTATATTTGGCATTCGAGTCTAAAG	1470
Db	598	ATGGGGTACCCGTGTGGAAGAACCAACCACTATATTTGGCATTCGAGTCTAAAG	657
OY	1471	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCCA	1530
Db	658	CATATGATACAGAGTACATATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCCA	717
OY	1531	ACCCACAGAAGATGTGATTTGGTAAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA	1590
Db	718	ACCCACAGAAGATGTGATTTGGTAAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA	777
OY	1591	TGTGAGAACAATGATCAGAGATATATATGATTTATGGATCAAAAGCCTAAAGCCATGTG	1650
Db	778	TGTGAGAACAATGATCAGAGATATATATGATTTATGGATCAAAAGCCTAAAGCCATGTG	837
OY	1651	TAAATTAACCCCACTGTGTGTAGTTTAAAGTGCACATGATTTGAAGAATGATCTATA	1710

Db	838	TAAATAATACCCACCTGTGTAGTTTAAAGTGCAGTATTGGAAGATATATCTATATA	897
Qy	1711	CCAATAGTGTAGACGGGAGAAATGATTAATGAGAAAGAGAGATATAAAACGTCTTTC	1770
Db	898	CCAATAGTGTAGACGGGAGAAATGATTAATGAGAAAGAGAGATATAAAACGTCTTTC	957
Qy	1771	ATATCAGCACAAGCATAGAGATTAAGGTGCAGAAAGAAATATGATCTTTTATTAAC	1830
Db	958	ATATCAGCACAAGCATAGAGATTAAGGTGCAGAAAGAAATATGATCTTTTATTAAC	1017
Qy	1831	ATATAGTACCATAAGTAA-----TACCAGTATAGTGTGATTAAGTTGTAAACCTC	1884
Db	1018	ATATATATACCATAAGTAAATGATTAAGTACTACAGCTATACGTTACAAAGTTGTAAACCTC	1077
Qy	1885	TCATTCACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCCAATTCACATCATATTTGTGCC	1944
Db	1078	TCATTCACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCCAATTCACATCATATTTGTGCC	1137
Qy	1945	CGCGTGGTTTGGCGATTCTTAAATGTAAATTAAGACCTTCATAGGAACAGACCATGTA	2004
Db	1138	CGCGTGGTTTGGCGATTCTTAAATGTAAATTAAGACCTTCATAGGAACAGACCATGTA	1197
Qy	2005	CAATTCACAGCAGTACATGTAACATGGAATGAGCCAGTATGATTCATCACTCACTGC	2064
Db	1198	CAATTCACAGCAGTACATGTAACATGGAATGAGCCAGTATGATTCATCACTCACTGC	1257
Qy	2065	TGTTTAATAGTGCAGTGTAGCAGAAAGATGTGTAAATAGATTCGCCAATTTTCACAGCA	2124
Db	1258	TGTTTAATAGTGCAGTGTAGCAGAAAGATGTGTAAATAGATTCGCCAATTTTCACAGCA	1317
Qy	2125	ATGCTTAAACCATATATAGTACAGCTGCAACACATCTGTAGAAATTAATTTGTACAAGCCCA	2184
Db	1318	ATGCTTAAACCATATATAGTACAGCTGCAACACATCTGTAGAAATTAATTTGTACAAGCCCA	1377
Qy	2185	ACAACATATCAGAAAAAATATCCGTTATCCAGAGGGGACCAAGGAGACATTTGTACAA	2244
Db	1378	ACAACATATCAGAAAAAATATCCGTTATCCAGAGGGGACCAAGGAGACATTTGTACAA	1437
Qy	2245	TAGCAAAAAATAGAAATATGAGACAAGCATTGTAAATTAAGTATGAGCAAAAAATGATG	2304
Db	1438	TAGCAAAAAATAGAAATATGAGACAAGCATTGTAAATTAAGTATGAGCAAAAAATGATG	1497
Qy	2305	CCACTTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATATTAATAACAATTA	2364
Db	1498	ACACTTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATATTAATAACAATTA	1557
Qy	2365	TCCTTTAAGCAATCCTAGAGAGGGGACCCAGAAATTTGTAACGCACAGTTTAAATGTGGAG	2424
Db	1558	TCCTTTAAGCAATCCTAGAGAGGGGACCCAGAAATTTGTAACGCACAGTTTAAATGTGGAG	1617
Qy	2425	GGGAATTTTCTACTGTAATTCACACAACTGTTTAAATAGTACTGGTTTAAATAGTACTT	2484
Db	1618	GGGAATTTTCTACTGTAATTCACACAACTGTTTAAATAGTACTGGTTTAAATAGTACTT	1677
Qy	2485	GGAGTACTGAGGGCTCAAAATAACACTGGAAGAAAGTACACATCACTCCATCGCAGAA	2544
Db	1678	GGAGTACTGAGGGCTCAAAATAACACTGGAAGAAAGTACACATCACTCCATCGCAGAA	1737
Qy	2545	TAAACCAATTTATTAACATGTGGCAGGAAGTATGGAAGAAAGCAATGTATGCCCTCCATCA	2604
Db	1738	TAAACCAATTTATTAACATGTGGCAGGAAGTATGGAAGAAAGCAATGTATGCCCTCCATCA	1797
Qy	2605	GTCGACCAATTTAGATGTCATCAAAATTTACTGGGCTCTATTAAACAAGATAGTGGTGA	2664
Db	1798	GTCGACCAATTTAGATGTCATCAAAATTTACTGGGCTCTATTAAACAAGATAGTGGTGA	1857
Qy	2665	ATTAACAACAATGGGTCCGAGATCTTTCAGACCTGGAGAGAGGAGATTTGAGGACAATTGGA	2724
Db	1858	ATTAACAACAATGGGTCCGAGATCTTTCAGACCTGGAGAGAGGAGATTTGAGGACAATTGGA	1917
Qy	2725	GAAATGAATTTATTAATATTAAGTATTAATAATTTGAACATTAAGAGATGACACCAACA	2784

Db	1918	AAAGTGATTTATTAATAATTAAGTACGTAATAAATTTGAACCTTATAGGATACGACCCACCA	1977
OY	2785	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATAGGAGCTTGTGTC	2844
Db	1978	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATAGGAGCTTGTGTC	2037
OY	2845	TTGGGCTCTGGGAGCAGCAGGAAGCACTATGGGCTGCAGCTCAATGACGCTGACGGTAC	2904
Db	2038	TTGGGCTCTGGGAGCAGCAGGAAGCACTATGGGCTGCAGCTCAATGACGCTGACGGTAC	2097
OY	2905	AGGCGACAGCAATTTATGTCTGATATAGTGCAGCAGCAGAAACAATTTGGTGGGCTATGTG	2964
Db	2098	AGGCGACAGCAATTTATGTCTGATATAGTGCAGCAGCAGAAACAATTTGGTGGGCTATGTG	2157
OY	2965	AGGCGCAACAGCATCTGTGGCACTCAGTCGTGGGGCATCAACAGCTCCAGGCCAAGAA	3024
Db	2158	AGGCGCAACAGCATCTGTGGCACTCAGTCGTGGGGCATCAACAGCTCCAGGCCAAGAA	2217
OY	3025	TTCCGTGGCTGGGAAAGATACCTAAAGATCAACAGCTCTGCGGATTTGGGGTTGCTGTG	3084
Db	2218	TTCCGTGGCTGGGAAAGATACCTAAAGATCAACAGCTCTGCGGATTTGGGGTTGCTGTG	2277
OY	3085	GAAGAACTATTTGGACCACTGCTGCGCTTGGATGCGTAGTTGGAGTAATTAATCTGCGG	3144
Db	2278	GAAGAACTATTTGGACCACTGCTGCGCTTGGATGCGTAGTTGGAGTAATTAATCTGCGG	2337
OY	3145	AACAGATTTGGAAATTAACATGACCTGGATGAGTGGGAGACAGAAATTAACATTAACAA	3204
Db	2338	AACAGATTTGGAAATTAACATGACCTGGATGAGTGGGAGAGAAATTAACATTAACAA	2397
OY	3205	GCTTAATTAACATCTCTTAATTGAATAATCGCAAAACGCAAGAAAGATGACAGAAAT	3264
Db	2398	GCTTAATTAACATCTCTTAATTGAATAATCGCAAAACGCAAGAAAGATGACAGAAAT	2457
OY	3265	TATTGGAATTTGATTAATGGGCAAGTTTGTGGAATTTGGTTAACATTAACAAATGGCGTG	3324
Db	2458	TATTGGAATTTGATTAATGGGCAAGTTTGTGGAATTTGGTTAACATTAACAAATGGCGTG	2517
OY	3325	GGATATATAAAATTAATCATTAATAGTAGAGGAGCTTGTAGCTTTAAGATATGTTTGG	3384
Db	2518	GGATATATAAAATTAATCATTAATAGTAGAGGAGCTTGTGTAGCTTTAAGATATGTTTGG	2577
OY	3385	CTGTACTTTTCTATAGTGAAATAGATTTAGGCAAGGATATCAACCTTATCGTTTAGACC	3444
Db	2578	CTGTACTTTTCTATAGTGAAATAGATTTAGGCAAGGATATCAACCTTATCGTTTAGACC	2637
OY	3445	ACCTCCCAATCCGACGAGGAGCCGACAGCGGCCGCAAGGAATTTGAAGAAGAGGTGGAGAA	3504
Db	2638	ACCTCCCAATCCGACGAGGAGCCGACAGCGGCCGCAAGGAATTTGAAGAAGAGGTGGAGAA	2697
OY	3505	GAGACAGAGACAGATTCATTCGATTAGTAGAAGGATCCTTAGCACTTATCTGGAGCAGTC	3564
Db	2698	GAGACAGAGACAGATTCATTCGATTAGTAGAAGGATCCTTAGCACTTATCTGGAGCAGTC	2757
OY	3565	TGCGGAGGCTGTGGCTCTTCAAGCTACCAACCGCTTGAAGACTTACTGTTGTTAAGCA	3624
Db	2758	TGCGGAGGCTGTGGCTCTTCAAGCTACCAACCGCTTGAAGACTTACTGTTGTTAAGCA	2817
OY	3625	GGATTTGGGAACTTCTGGAGCGCAGGGGGGTGGGAAGGCGTCAAAATTTGGGGAATCTCC	3684
Db	2818	GGATTTGGGAACTTCTGGAGCGCAGGGGGGTGGGAAGGCGTCAAAATTTGGGGAATCTCC	2877
OY	3685	TACAGTATTTGGAGTCAGGAATTAAGAAATAGTGTGTTAACTTGCTCAATGCCACAGCA	3744
Db	2878	TACAGTATTTGGAGTCAGGAATTAAGAAATAGTGTGTTAACTTGCTCAATGCCACAGCTA	2937
OY	3745	TACAGTAGCTGAGGGGAGACAGATAGGGTTATAGAAGTATTAACAAGCTTATAGAGCTA	3804
Db	2938	TACAGTAGCTGAGGGGAGACAGATAGGGTTATAGAAGTATTAACAAGCTTATAGAGCTA	2997
OY	3805	TTGCCCACTAATCTAGAAAGATTAAGACAAGGCTTGGAAAGGATTTTCTATAGATGGGT	3864
Db	2998	TTGCCCACTAATCTAGAAAGATTAAGACAAGGCTTGGAAAGGATTTTCTATAGATGGGT	3057

```

RESULT 9
US-10-003-035-14
; Sequence 14, Application US/10003035
; Publication NO. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 2248-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 3157
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Env/Tat/Rev
US-10-003-035-14

```

Query Match	41.68; Score 2592.4; DB 13; Length 3157;
-------------	--

Matches 2633; Conservative 0; Mismatches 41; Indels 6; Gaps 1;

QY 1291 CCATGAGAGTGAAGGAGAAGTATCAGCACTTGTGGAGATGGGGGTGGAATGGGCACCA 1350

Db 478 CAATGAGACTGAAGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAGATGGGGCACCA 537

1351 TGGCTCCCTGGGATATTGATGATCTGTAGTGCACAGAAAAATTGTGGGTCAACGCTATT 1410

DB 538 1GCLCC1GGGAG1GTGA1GAC1G1AG1GCTACAGAAAAT1G1GGG1CACAG1CIAIT 591

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

1531 ACCCACAAGTATTTGGTAAATGTGACAGAAATTTTAACTATGTGGAAATGACA 1590

Db 718 ACCACAGAGTACTATTGCTAAATCTGCACAGAAATTTTAAACATCTGGAAAAATGACA 777

QY 1591 TGGTAGACAGATGCATGAGGATATAATCAGTTTATGGGATCAAGCCTAAGCCATGTG 1650

Db 778 TGGTAGAACAGATGCATGAGGATATATCAGTTATGGGATCAAAGCCCTAAAGCCATGTG 837

2Y 1651 TAAATTAACCCCACTCTGTAGTTTAAAGTGCAC TGATTGAGATGATACTATA 1710

Db 838 TAAATTTAACCCCACTCTGTGTTAGTTTAAAGTGCACCTGATTTGAAGATGATACTATA 897

1/11 CCAATAGTAGTAGCGGAGAAATGATTAATGGAGAAAGGAGAGATAAAAACCTGCTTTCA 1/70

070 CCAAGTAGTAGCCTGGGAAATGAATAACCGAAGAAGTATTACTTCATGCCTTTCCATGTG

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	



1831 ATATAGTACCAATATGATA-----TACACAGCTATAGTTGATAGTTGTAACACGCTAG 1884  
1018 ATATTAATACCAATATGATATGATATCTACACAGCTATAGTTGATACAGTTGTACACCTCAG 1077  
1885 TCATTACACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCCTACATTAATTTGTGCC 1944  
1078 TCATTACACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCCTACATTAATTTGTGCC 1137  
1945 CGGCTGGTTTGGCATTTCTAAAATGTATATAGAGCTTCAATGGAACAGGACCATGTA 2004  
1138 CGGCTGGTTTGGCATTTCTAAAATGTATATAGAGCTTCAATGGAACAGGACCATGTA 1197  
2005 CAAATGTACAGCAGTACATATGTACATGGAATCAGGACAGTATGATCAATCACTGAC 2064  
1198 CAAATGTACAGCAGTACATATGTACATGGAATTAGGCCAGTATGATCAATCACTGAC 1237  
2065 TGTAAATGGCAGCTGTAGCAGAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCA 2124  
1258 TGTAAATGGCAGCTGTAGCAGAGAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCA 1317  
2125 ATGCTAAACCATATATAGTACAGTGAACACATCTGTAGAAATTAATTTGTACAGACCCA 2184  
1318 ATGCTAAACCATATATAGTACAGTGAACACATCTGTAGAAATTAATTTGTACAGACCCA 1377  
2185 ACAACATACAGAGAGAGAGATGTCGATCCAGAGGGAGCCAGGAGAGCATTTGTACAA 2244  
1378 ACAACATACAGAGAGAGAGATGTCGATCCAGAGAGAGCCAGGAGAGCATTTGTACAA 1437  
2245 TAGGAGAGAGATGAGAAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATG 2304  
1438 TAGGAGAGAGATGAGAAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATA 1497  
2305 CCACCTTTAAACAGATAGCTATAGCAAAATTAAGAGACAAATTTGGAATTAATTAACATTA 2384  
1498 ACACCTTTAAACAGATAGCTATAGCAAAATTAAGAGACAAATTTGGAATTAATTAACATTA 1557  
2365 TCCTTAAGCAATCCTCAGAGAGGGAGCCAGAAATTTGTACGACAGTATTTAATTTGTGAG 2424  
1558 TCCTTAAGCAATCCTCAGAGAGGGAGCCAGAAATTTGTACGACAGTATTTAATTTGTGAG 1617  
2425 GGGAAATTTTCTACTGTATTAATTCACACAACTGTTTAAATAGTACTGTTTAAATAGTACT 2484  
1618 GGGAAATTTTCTACTGTATTAATTCACACAACTGTTTAAATAGTACTGTTTAAATAGTACT 1677  
2485 GGAATCTAGAGAGGTCAATTAACACTGAAGAAATGACACATCTCCATGCAAGAA 2544  
1678 GGAATCTAGAGAGGTCAATTAACACTGAAGAAATGACACATCTCCATGCAAGAA 1737  
2545 TAAACAAATTTATTAACATGTGGCAGAGAGTAAAGAAAGCAATATGAGCCCTCCATCA 2604  
1738 TAAACAAATTTATTAACATGTGGCAGAGAGTAAAGAAAGCAATATGAGCCCTCCATCA 1797  
2605 GTGAGACAAATTAGATGTTCAATTAATTTACTGGGCTGTATTAACAGAGATGTTGTTA 2664  
1798 GTGAGACAAATTAGATGTTCAATTAATTTACTGGGCTGTATTAACAGAGATGTTGTTA 1857  
2665 ATTAACAACATGGTCCGAGATCTTCAGAGCTGGAGAGGCGATATGAGGACAAATTTGA 2724  
1858 ATTAACAACATGGTCCGAGATCTTCAGAGCTGGAGAGGCGATATGAGGACAAATTTGA 1917  
2725 GAAATGATTTATTAATTAATTAATTAATTTGAACCTTTAGAGGTAGCACCCAGCA 2784  
1918 GAAATGATTTATTAATTAATTAATTAATTTGAACCTTTAGAGGTAGCACCCAGCA 1977  
2785 AGGCAAGAGAGAGAGTGTGAG 2844  
1978 AGGCAAGAGAGAGAGTGTGAG 2037  
2845 TTGGGTTCTTGGAG 2904  
2038 TTGGGTTCTTGGAG 2097

2905 AGGCAAGAGAGAGAGTGTGAG 2964  
2998 AGGCAAGAGAGAGTGTGAG 2157  
2965 AGGCAAGAGAGAGTGTGAG 3024  
2158 AGGCAAGAGAGAGTGTGAG 2217  
3025 TCCGAGCTGGAGAGAGATTAAGATCAACAGCTTCGAGGATTTGGGGTGTGCTG 3084  
2218 TCCGAGCTGGAGAGAGATTAAGATCAACAGCTTCGAGGATTTGGGGTGTGCTG 2277  
3085 GAAACATCAATTTGACACACTGCTGAGCTTGGAGATGCTAGTGGAGTAAATTAATCTGCG 3144  
2278 GAAACATCAATTTGACACACTGCTGAGCTTGGAGATGCTAGTGGAGTAAATTAATCTGCG 2337  
3145 AACAGATTTGGAATTAACATGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3204  
2338 AACAGATTTGGAATTAACATGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397  
3205 GCTTAATACACTGCTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3264  
2398 GCTTAATACACTGCTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457  
3265 TATTGGAATTAAGATTAATGAG 3324  
2458 TATTGGAATTAAGATTAATGAG 2517  
3325 GGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3384  
2518 GGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2577  
3385 CTGATCTTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3444  
2578 CTGATCTTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2637  
3445 ACCTCCCAATCCCGAGAGGAGCCGACAGAGCCGGAAGAAATTAAGAGAGAGAGAGAGAGAG 3504  
2638 ACCTCCCAATCCCGAGAGGAGCCGACAGAGCCGGAAGAAATTAAGAGAGAGAGAGAGAGAG 2697  
3505 GAGACAGAGACAGATTCATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3564  
2698 GAGACAGAGACAGATTCATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2757  
3565 TGGGAGAGCTGTGCTCTTCAAGTACCAACCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3624  
2758 TGGGAGAGCTGTGCTCTTCAAGTACCAACCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2817  
3625 GGATTTGGAACCTTCTGGAG 3684  
2818 GGATTTGGAACCTTCTGGAG 2877  
3685 TACAGTATTTGAGAGTACAGAACTAAAGAAATAGTGTCTTAACTTCTCAATGTCACAGCA 3744  
2878 TACAGTATTTGAGAGTACAGAACTAAAGAAATAGTGTCTTAACTTCTCAATGTCACAGCA 2937  
3745 TAGCAGTAGCTGAGAGGAGAGATAGGTTATTAAGAGATTAACAGCAGCTTATAGAGCTA 3804  
2938 TAGCAGTAGCTGAGAGGAGAGATAGGTTATTAAGAGATTAACAGCAGCTTATAGAGCTA 2997  
3805 TTGCGCACATTAATTAAG 3864  
2998 TTGCGCACATTAATTAAG 3057  
3865 GGCAGAGTGTCAAAAGTAGTGTATTTGATGGCTGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 3924  
3058 GGCAGAGTGTCAAAAGTAGTGTATTTGATGGCTGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 3117  
3925 GCTGAGCCAGCAGCAGATGGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3964  
3118 GCTGAGCCAGCAGCAGATGGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3157

```
RESULT 10
; US-10-286-332A-22
; Sequence 22, Application US/10286332A
; Publication No. US20030138459A1
; GENERAL INFORMATION:
; APPLICANT: Manu, Danher
; TITLE OF INVENTION: METHOD OF VACCINATION THROUGH SEROTYPE ROTATION
; FILE REFERENCE: 22488-738
; CURRENT APPLICATION NUMBER: US/10/286,332A
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 10/003,035
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/NeF from strain BH10
US-10-286-332A-22

Query Match      41.4%; Score 2576; DB 12; Length 2950;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 45; Indels 12; Gaps 2;

QY      1286  CGCCACCATGAGATGAAAGGAAGATCAGCACTTGTGAGATGGGGGTGGAATGGG 1345
DB      6      CGCCACCATGGAGTGAAGGAATATCAGCACTTGTGAGATGGGGGTGGAATGGG 65

QY      1346  CACCATCTCTCTGGATATGATGATCTGTAGTGTCTACAGAAAAATTTGGTGCACCGT 1405
DB      6      CACCATCTCTCTGGATATGATGATCTGTAGTGTCTACAGAAAAATTTGGTGCACCGT 125

QY      1406  CTATATATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTTTGTGCATCAGATGC 1465
DB      126     CTATATATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTTTGTGCATCAGATGC 185

QY      1466  TAAAGCATATGATACAGAGTACATATATGTTGGGCCACACATCCCTGTGTACCCACAGA 1525
DB      186     TAAAGCATATGATACAGAGTACATATATGTTGGGCCACACATCCCTGTGTACCCACAGA 245

QY      1526  CCCCACCCACAGAAGATGATGTAATGTGACAGAAAAATTTTAAACATGTGAAAAA 1585
DB      246     CCCCACCCACAGAAGATGATGTAATGTGACAGAAAAATTTTAAACATGTGAAAAA 305

QY      1586  TGACATGATAGAACAGATGATGATGATATATCAGTTTATGGGATCAAAAGCTTAAAGCC 1645
DB      306     TGACATGATAGAACAGATGATGATGATATATCAGTTTATGGGATCAAAAGCTTAAAGCC 365

QY      1646  ATGCTGTAATTAACCCACATCTGTGTAGTTTAAAGTGCATGTTGAAGATGATAC 1705
DB      366     ATGCTGTAATTAACCCACATCTGTGTAGTTTAAAGTGCATGTTGAAGATGATAC 425

QY      1706  TAATACCAATAGTATGAGCGGAGATGATATATGAGAGAAAGAGATTAATAAACTGCTC 1765
DB      426     TAATACCAATAGTATGAGCGGAGATGATATATGAGAGAAAGAGATTAATAAACTGCTC 485

QY      1766  TTTCATATATGACACAGCATATAGATTAAGGTGCAGAAAGAAATATGATCTTTTATAA 1825
DB      486     TTTCATATATGACACAGCATATAGATTAAGGTGCAGAAAGAAATATGATCTTTTATAA 545

QY      1826  ACTTGATATATAGTACCATATAGTAA-----TACCAAGCTATATGTTGATAGTTAAAC 1879
DB      546     ACTTGATATATATACCAATATAGTATATGATACCAAGCTATATGTTGATAGTTAAAC 605

QY      1880  CTCAGTATTTACAGAGCGCTGTCCAAAGTATCTCTTGGAGCAATTCCTACATATATG 1939
DB      606     CTCAGTATTTACAGAGCGCTGTCCAAAGTATCTCTTGGAGCAATTCCTACATATATG 665
```

```
QY      1940  TGCCCGCGCTGTTTGCATATCTTAAATGTAATATAGACGTTCAATGGAACAGGACC 1999
DB      666     TGCCCGCGCTGTTTGCATATCTTAAATGTAATATAGACGTTCAATGGAACAGGACC 725

QY      2000  ATGTACAATGTGACACAGTACATATGATGATGAAATGAGCCAGTATATCACTCA 2059
DB      726     ATGTACAATATGTGACACAGTACATATGATGATGAAATGAGCCAGTATATCACTCA 785

QY      2060  ACTGCTGTAAATGCGAGTGTAGCAGAAAGATGTATATATGATCTGCCAATTTTAC 2119
DB      786     ACTGCTGTAAATGCGAGTGTGCGAAGAAAGATGTATATGATCTGCCAATTTTAC 845

QY      2120  AGACAATGCTAAAAACCTAATATGATACGCTGAACACATCTGTGAAATTAATGTACA 2179
DB      846     AGACAATGCTAAAAACCTAATATGATACGCTGAACACATCTGTGAAATTAATGTACA 905

QY      2180  ACCCAACACATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGCATTTGT 2239
DB      906     ACCCAACACATACAGAAAAAGTATCCGTATCCAGAGAGGACCGAGAGCATTTGT 965

QY      2240  TACATATGAAAAATAGAAATATGACAAAGCATTTGTATACATTTAGTACGCAAAATG 2299
DB      966     TACATATGAAAAATAGAAATATGACAAAGCATTTGTATACATTTAGTACGCAAAATG 1025

QY      2300  GAATGCCACTTTAAACAGATAGCTACCAATTAAGAGAACAAATTTGAAATATATAAC 2359
DB      1026     GAATTAACACTTTAAACAGATAGCTACCAATTAAGAGAACAAATTTGAAATATATAAC 1085

QY      2360  AATATCTTTTAAGCAATCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATG 2419
DB      1086     AATATCTTTTAAGCAATCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATG 1145

QY      2420  TGGAGGGGAATTTTCTACGTATATCAACAACTGTTTAATATGCTGTGTTAATAG 2479
DB      1146     TGGAGGGGAATTTTCTACGTATATCAACAACTGTTTAATATGCTGTGTTAATAG 1205

QY      2480  TACTTGAATAGTGAAGGGTCAATAAACACTGAAGGAAGTACAAATCACTCCCATG 2539
DB      1206     TACTTGAATAGTGAAGGGTCAATAAACACTGAAGGAAGTACAAATCACTCCCATG 1265

QY      2540  CAGATTAACAAANTTTAAACNTGTGGCAGGAAGTAGGAAGAACATATGATGCCCTCC 2599
DB      1266     CAGATTAACAAANTTTAAACATGTGGCAGGAAGTAGGAAGAACATATGATGCCCTCC 1325

QY      2600  CATCAGTGAACAATTTGATGTTTCATCAATATATACAGGCTGTATTAACAAGATAG 2659
DB      1326     CATCAGTGAACAATTTGATGTTTCATCAATATATACAGGCTGTATTAACAAGATAG 1385

QY      2660  TGCTAATTAACAACATGGGTCCGAGATCTTCAGACCTGAGAGGCGATATGAGGACAA 2719
DB      1386     TGCTAATTAACAACATGATGCCGAGATCTTCAGACCTGAGAGGAGATATGAGGACAA 1445

QY      2720  TTGGAGAAGTGAATTAATTAATTAAGTAAATTTGAACCATTTAGAGTAGACACC 2779
DB      1446     TTGGAGAAGTGAATTAATTAATTAAGTAAATTTGAACCATTTAGAGTAGACACC 1505

QY      2780  CACCAAGCGCAAGAGAGAGTGTGACAGAGAAAAAGAGACAGTGGGAATAGAGCTTT 2839
DB      1506     CACCAAGCGCAAGAGAGAGTGTGACAGAGAAAAAGAGACAGTGGGAATAGAGCTTT 1559

QY      2840  GTTCTTGGGTTCTTTGGAGCAGCAGAACCATATGGGCTGCACGTCAATGACGCTGAC 2899
DB      1560     GTTCTTGGGTTCTTTGGAGCAGCAGAACCATATGGGCGCAGCGTCATGACGCTGAC 1619

QY      2900  GGTACAGGCCACAGACATATATGTCTGATATATGTCACACACAGACAAATTTGCTG 2959
DB      1620     GGTACAGGCCACAGACATATATGTCTGATATATGTCACACACAGACAAATTTGCTG 1679

QY      2960  TATGAGGGCGACAGCATCTGTGTGCAACGTACAGTCTGGGGCATCAAGAGCTCCAGGC 3019
DB      1680     TATGAGGGCGACAGCATCTGTGTGCAACGTCTGGGGCATCAAGAGCTCCAGGC 1739
```

QY	3020	AAGAAATCTGGCTGTGGAAAGATATCCCTAAAGATCAACAGCTCCTGGGAAATTGGGGTGG	3079
Db	1740	AAGATATCTGGCTGTGGAAAGATATCCCTAAAGATCAACAGCTCCTGGGAAATTGGGGTGG	1799
QY	3080	CTCTGAGAAATCTATTTGGACACACAGCTGCTGCTCTTGGAAATCTGATGGAGTAATTAATC	3139
Db	1800	CTCTGAGAAATCTATTTGGACACACAGCTGCTGCTCTTGGAAATCTGATGGAGTAATTAATC	1859
QY	3140	TCCTGGAACAGATTTGGAAATACATGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA	3199
Db	1860	TCCTGGAACAGATTTGGAAATACATGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA	1919
QY	3200	CACAACTTATACACTTCCTTAATTTGAAGATTCGCAAAACGACAGAGAAAAGATGAAACA	3259
Db	1920	CACAACTTATACACTTCCTTAATTTGAAGATTCGCAAAACGACAGAGAAAAGATGAAACA	1979
QY	3260	AGAAATTTATGAAATAGATTAATGGGACAGTTTGGAAATGGTTTAACATACAAATTTG	3319
Db	1980	AGAAATTTATGAAATAGATTAATGGGACAGTTTGGAAATGGTTTAACATACAAATTTG	2039
QY	3320	GCTGTGGTATATATAAATTTATTCATATATGATAGTAGAGAGCTTGGTAGTGTTAAGAAATAGT	3379
Db	2040	GCTGTGGTATATATAAATTTATTCATATATGATAGTAGAGAGCTTGGTAGTGTTAAGAAATAGT	2099
QY	3380	TTTTGGCTGTAATTTCTATATGAAATAGAGTTAGCCAGGAGATATTCACCATTTATCGTTTCA	3439
Db	2100	TTTTGGCTGTAATTTCTATATGAAATAGAGTTAGCCAGGAGATATTCACCATTTATCGTTTCA	2159
QY	3440	GACCCACTCCCATCCCGAGGGGACCCGACAGGCCGGAAGAGATTAAGAAAGAAAGTGG	3499
Db	2160	GACCCACTCCCATCCCGAGGGGACCCGACAGGCCGGAAGAGATTAAGAAAGAAAGTGG	2219
QY	3500	AGAGAGAGACAGAGACAGATCCATTCGATTAAGTAAACGATCCTTAGCATTATCTGGGA	3559
Db	2220	AGAGAGAGACAGAGACAGATCCATTCGATTAAGTAAACGATCCTTAGCATTATCTGGGA	2279
QY	3560	CGATCTGCGGAGCTGTGCTCTTCACGTACACACCGCTTGAGAGACTTACTCTTGATTTGT	3619
Db	2280	CGATCTGCGGAGCTGTGCTCTTCACGTACACACCGCTTGAGAGACTTACTCTTGATTTGT	2339
QY	3620	AACGAGAGATTTGGAACTTCTGGGACGAGGGGGTGGGAAGCCCTCAAAATTTGGTGAA	3679
Db	2340	AACGAGAGATTTGGAACTTCTGGGACGAGGGGGTGGGAAGCCCTCAAAATTTGGTGAA	2399
QY	3680	TCCTCTCAGTATTTGGAGTCAAGAACTAAAGAAATAGTGTCTTAACTTGTCTCAATGCCAC	3739
Db	2400	TCCTCTCAGTATTTGGAGTCAAGAACTAAAGAAATAGTGTCTTAACTTGTCTCAATGCCAC	2459
QY	3740	AGCCATAGCAGTACTGAGGGGACAGATAGGGTTATAGAAATATTAACAGCAGCTTATAG	3799
Db	2460	AGCCATAGCAGTACTGAGGGGACAGATAGGGTTATAGAAATATTAACAGCAGCTTATAG	2519
QY	3800	AGCATTTGGCCACTACTAGAAACAAATTAAGACAGGCTTGGAAAGGATTTTGGCTATAGA	3859
Db	2520	AGCATTTGGCCACTACTAGAAACAAATTAAGACAGGCTTGGAAAGGATTTTGGCTATAGA	2579
QY	3860	TGGGTGGCAAGTGTCTCAAAAAGTATGATGTGATGGCTGCTGTAAAGGAAAGAAATGA	3919
Db	2580	TGGGTGGCAAGTGTCTCAAAAAGTATGATGTGATGGCTGCTGTAAAGGAAAGAAATGA	2639
QY	3920	GACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAATCTTAGA 3971	
Db	2640	GACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGATCTCGAAGACTTAGA 2691	

/	CURRENT APPLICATION NUMBER:	US/10/003_035
/	CURRENT FILING DATE:	2001-11-01
/	PRIOR APPLICATION NUMBER:	09/585_599
/	PRIOR FILING DATE:	2000-06-02
/	PRIOR APPLICATION NUMBER:	PCT/US01/18238
/	PRIOR FILING DATE:	2001-06-04
/	NUMBER OF SEQ ID NOS:	75
/	SOFTWARE:	PatentIn version 3.1
/	SEQ ID NO 22	
/	LENGTH:	2950
/	TYPE:	DNA
/	ORGANISM:	Artificial sequence
/	FEATURE:	
/	OTHER INFORMATION:	Modified Env/Nef from strain BH10
/	US-10-003-035-22	
Query Match	41.4%;	Score 2576; DB 13; Length 2950;
Best Local Similarity	97.9%;	Pred. No. 0;
Matches 2635;	Conservative	0; Mismatches 45; Indels 12; Gaps 2;
OY	1286	CGCCACCATGAGATGAAGAATAATCAGACTTGTGGAGATGGGGTGAAATGGG 1345
Db	6	CGCCACCATGGAGAGTAAGAAATATCAGACACTTGTGGAGATGGGGTGAGATGGG 65
OY	1346	CACCATGCTCTTTGGGATATTGANGATCTAGTCCTACAGAAAAATTGTGGTCACCGT 1405
Db	66	CACCATGCTCTTTGGGATATTGATGATCTGATGTCCTACAGAAAAATTGTGGTCACAGT 125
OY	1406	CTATTATGGGGTACTGTGTGGAAGAACCAACCACTCTATTTTGTGCATCAGATGC 1465
Db	126	CTATTATGGGGTACTGTGTGGAAGAACCAACCACTCTATTTTGTGCATCAGATGC 185
OY	1466	TAAACATATGATACAGAGATACATAAATGTTTTGGGCCACACATGCGCTGTACCCACAGA 1525
Db	186	TAAACATATGATACAGAGATACATAAATGTTTTGGGCCACACATGCGCTGTACCCACAGA 245
OY	1526	CCCCAACCCACAGAAGTAGTATTGGTAAATGTGCACAGAAAATTTTAAACATGTGGAAAA 1585
Db	246	CCCCAACCCACAGAAGTAGTATTGGTAAATGTGCACAGAAAATTTTAAACATGTGGAAAA 305
OY	1586	TGACATGTGTAGAACGATGCATGAGGATATATCATGTTTATNGGATCAAAACCTTAAAGCC 1645
Db	306	TGACATGTGTAGAACGATGCATGAGGATATATCATGTTTATNGGATCAAAACCTTAAAGCC 365
OY	1646	ATGTSTAAATTTAAACCCCACCTCTGTGTAGTTTAAAGTGCACCTGTTTGAAGAAATGATAC 1705
Db	366	ATGTSTAAATTTAAACCCCACCTCTGTGTAGTTTAAAGTGCACCTGTTTGAAGAAATGATAC 425
OY	1706	TAATACCAATAGTAGTAGCGGGAGAAATGATATATGGAGAAAGAGATAAAAAACTGCTC 1765
Db	426	TAATACCAATAGTAGTAGCGGGAGAAATGATATATGGAGAAAGAGATAAAAAACTGCTC 485
OY	1766	TTTCAATATFCAGACACAACATTAAGATTAAGTGCAGAAAGAAATATGCAITCTTTTATTA 1825
Db	486	TTTCAATATFCAGACACAACATTAAGATTAAGTGCAGAAAGAAATATGCAITCTTTTATTA 545
OY	1826	ACTTGATTTAGTACCAATATAGATTA-----TACACGCTATATGTTGATATAGCTTGTAAAC 1879
Db	546	ACTTGATTTAGTACCAATATAGATTAAGTACTACACGCTATATGCTTGTGACAAAGTTGTAAAC 605
OY	1880	CTCAGTCATTAACACAGGCGCTGCCAAAGGATATCCTTTAGGCCAATTCOCATACATTAATG 1939
Db	606	CTCAGTCATTAACACAGGCGCTGCCAAAGGATATCCTTTAGGCCAATTCOCATACATTAATG 665
OY	1940	TGCCCCGGCTGTGTTTGGGATTTCTTAAATGTATATATATTAAGCGTTTCAATGGAACAGACC 1999
Db	666	TGCCCCGGCTGTGTTTGGGATTTCTTAAATGTATATATTAAGCGTTTCAATGGAACAGACC 725
OY	2000	ATGTCAATATGTCAGACAGATACATATGATACATGGAATTCAGGCGAGATGATCAACATCA 2059
Db	726	ATGTCAATATGTCAGACAGATACATATGATACATGGAATTTAGGCGAGATGATCAACTCA 785
OY	2060	ACTGCTGTAAATGGCAGCTTACGAGAAGATGTAGTAATTTAGATCTGCCAATTTTCCAC 2119





DB	Sequence	DB 12	Length	Score	DB 12	Length	Score
DB	8142 GGTATATATAATATTCATTAATGATAGTAGGGCTGTGGATTGAAGATAGCTTTTG	8201	100	41.3%	8201	100	2571
OY	3385 CTGTACTTTCTATAGTGAATAGAGTTAGGCGAGGGATATTCACCAATATACGTTTGACACC	3444	100	97.9%	3444	100	2571
DB	8202 CTGTACTTTCTATAGTGAATAGAGTTAGGCGAGGGATATTCACCAATATACGTTTGACACC	8261	100	0	8261	100	2571
OY	3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGAAATAGAAAGAGAGGTGGAGAGA	3504	100	45	3504	100	2571
DB	8262 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGAAATAGAAAGAGAGGTGGAGAGA	8321	100	12	8321	100	2571
OY	3505 GAGACAGAGACAGATCCATTCGATTAGGAAGCATTCCTTAGCATTATCTGGGACGATC	3564	100	0	3564	100	2571
DB	8322 GAGACAGAGACAGATCCATTCGATTAGGAAGCATTCCTTAGCATTATCTGGGACGATC	8381	100	0	8381	100	2571
OY	3565 TCGGAGGCTGTGCTCTTCACGCTACACCCGTTGAGAGACTTACTCTTGGATTGTAACGA	3624	100	0	3624	100	2571
DB	8382 TCGGAGGCTGTGCTCTTCACGCTACACCCGTTGAGAGACTTACTCTTGGATTGTAACGA	8441	100	0	8441	100	2571
OY	3625 GGATTGTGGAACCTTCTGGGACGACAGGGGTGGAAAGCCCTCAATATTTGGTGAATCC	3684	100	0	3684	100	2571
DB	8442 GGATTGTGGAACCTTCTGGGACGACAGGGGTGGAAAGCCCTCAATATTTGGTGAATCC	8501	100	0	8501	100	2571
OY	3685 TACACTATTGGAGTGTAGGAACCTAAAGATTAAGTATCTGCTCAATGCCACAGCA	3744	100	0	3744	100	2571
DB	8502 TACACTATTGGAGTGTAGGAACCTAAAGATTAAGTATCTGCTCAATGCCACAGCA	8561	100	0	8561	100	2571
OY	3745 TAGCAGTGTAGGAGGAGACAGATAGGGTTTATAGAAGTTTACCAAGACCTATAGAGCTA	3804	100	0	3804	100	2571
DB	8562 TAGCAGTGTAGGAGGAGACAGATAGGGTTTATAGAAGTTTACCAAGACCTATAGAGCTA	8621	100	0	8621	100	2571
OY	3805 TTCGCCACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGGATTTGCTATTAAGATGGT	3864	100	0	3864	100	2571
DB	8622 TTCGCCACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGGATTTGCTATTAAGATGGT	8681	100	0	8681	100	2571
OY	3865 GGCAGTGTGTCAAAAAGTAGTGTATTTGATTTGGCTGTGTAGGAAAGATAGACGA	3924	100	0	3924	100	2571
DB	8682 GGCAGTGTGTCAAAAAGTAGTGTATTTGATTTGGCTGTGTAGGAAAGATAGACGA	8741	100	0	8741	100	2571
OY	3925 GCTGAGCCAGACAGATGGGGTGGAGACATATCTGATGATCTAGCA	3971	100	0	3971	100	2571
DB	8742 GCTGAGCCAGACAGATGGGGTGGAGACATATCTGATGATCTAGCA	8788	100	0	8788	100	2571



QY	2365	TCCTTTAAGCAATCCTCCAGAGGAGGCCACGAAATTTGTAAGCCACAGTTTAAATGTGGAG	2423
Db	1558	TCCTTTAAGCAATCCTCCAGAGGAGGCCACGAAATTTGTAAGCCACAGTTTAAATGTGGAG	1617
QY	2425	GGCAATTTTCTTACTGTAAATCAACACACACAGTTTAAATAGTACTGTGTTAAATAGTACTT	2484
Db	1618	GGCAATTTTCTTACTGTAAATCAACACACACAGTTTAAATAGTACTGTGTTAAATAGTACTT	1677
QY	2485	GGAGTACTGAAGGGTCAAAATPAACACTGAAAGAAAGTGCACAAATCACAACACTCCCATGACAGAA	2544
Db	1678	GGAGTACTGAAGGGTCAAAATPAACACTGAAAGAAAGTGCACAAATCACAACACTCCCATGACAGAA	1737
QY	2545	TAAACAATTTATTAACATGTGGCAGAGAAAGTGAAGAAAAAGCAATGTATGGCCCTCCCATCA	2604
Db	1738	TAAACAATTTATTAACATGTGGCAGAGAAAGTGAAGAAAAAGCAATGTATGGCCCTCCCATCA	1797
QY	2605	GTGACAAATTTAGATGTTTCATCAATATATTACTGGGCTGCTATTAAACAAGATGGTGSTA	2664
Db	1798	GTGACAAATTTAGATGTTTCATCAATATATTACTGGGCTGCTATTAAACAAGATGGTGSTA	1855
QY	2665	ATTAACAACATGGGTCGAGATCTTCAGACCTGGAAGAGGCGATATGAGGACAAATTGGA	2724
Db	1858	ATTAACAACATGGGTCGAGATCTTCAGACCTGGAAGAGGCGATATGAGGACAAATTGGA	1911
QY	2725	GAACTGAAATTTATTAATATTAACATGATGAATAATGAACCAATTTAGATAGTACCACCCACCA	2784
Db	1918	GAACTGAAATTTATTAATATTAACATGATGAATAATGAACCAATTTAGATAGTACCACCCACCA	1977
QY	2785	AGGCCAAAGAGAAAGTGGTCGACAGAGAAAAAGCAAGTGGGAATAGACCTTGTGTTCC	2844
Db	1978	AGGCCAAAGAGAAAGTGGTCGACAGAGAAAAAGCAAGTGGGAATAGACCTTGTGTTCC	2033
QY	2845	TTGGGTTCTTGGAGACACAGAGAAAGCACTATGGGCTGCACGTCAATGACGCTGACGGTAC	2904
Db	2032	TTGGGTTCTTGGAGAGACAGAGAAAGCACTATGGGCGCGCCTCAATGACGCTGACGGTAC	2091
QY	2905	AGGCGACAGCAATTTGTCTGATATPATGAGCAGACAGAGAAACAAATTTGCTGAGGGCTATATG	2966
Db	2092	AGGCGACAGCAATTTGTCTGATATPATGAGCAGACAGCAACAAATTTGCTGAGGGCTATATG	2151
QY	2965	AGGCGCAACAGACATCTGTTGGCACTCACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA	3024
Db	2152	AGGCGCAACAGACATCTGTTGGCACTCACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA	2211
QY	3025	TCCTGGCTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTGCTCTG	3084
Db	2212	TCCTGGCTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTGCTCTG	2271
QY	3085	GAAACATCATTTGGACACATCTGCTGCTTGGAAATGCTATGTTGGAGTAAATATCTCTGG	3144
Db	2272	GAAACATCATTTGGACACATCTGCTGCTTGGAAATGCTATGTTGGAGTAAATATCTCTGG	2331
QY	3145	AACAGATTTTGGAAATPAACATGACCTGAGTGGAGTGGAGAGAGAAATTAACATTTACACAA	3204
Db	2332	AACAGATTTTGGAAATPAACATGACCTGAGTGGAGTGGAGAGAGAAATTAACATTTACACAA	2391
QY	3205	GCTTAAATACATCTCTAATTGGAAGATTCGCAAAACACAGCAAGAAAAAGATGAACAAGAT	3266
Db	2392	GCTTAAATACATCTCTAATTGGAAGATTCGCAAAACACAGCAAGAAAAAGATGAACAAGAT	2451
QY	3265	TATTGGAAATTAAGATTAATGGGCAAGTTTGTGGAATTTGGTTTAACATAACAAATTTGGCTGT	3324
Db	2452	TATTGGAAATTAAGATTAATGGGCAAGTTTGTGGAATTTGGTTTAACATAACAAATTTGGCTGT	2511
QY	3325	GGTATTAATAAATTTATCATATATGATGATGAGAGGCTGTGATGTTTAAAGATAGTATTG	3384
Db	2512	GGTATTAATAAATTTATCATATATGATGATGAGAGGCTGTGATGTTTAAAGATAGTATTG	2571
QY	3385	CTGTACTTTTCTATATAGTAATAGAGTTTGAAGCAGGATATTCACATTAATCGTTTCAGACCC	3444
Db	2572	CTGTACTTTTCTGTATGTAATAGATTTAGCAGGAGATATTCACATTAATCGTTTCAGACCC	2631
QY	3445	ACCTCCCAATCCGAGGAGGCCGACAGAGCCCGAAGAGATAGAAAGAAAGGTGGAGAGA	3504

Db	2632	ACCTCCCAATCCGAGGGGACCCGACAGGCCGAGAGATTAAGAGAGAGGTGGAGGA	2691
QY	3505	GAGCAGAGAGCAGATCCATTTCGATTAGTAAGAGATCCCTTACACTTATCTGGACGATC	3564
Db	2692	GAGCAGAGAGCAGATCCATTTCGATTAGTAAGAGATCCCTTACACTTATCTGGACGATC	2751
QY	3565	TGCGGAGCCCTGTCCTTCCTACGTACCCAGCCTTGAGAGACTTACTCTTGATTGAACGA	3624
Db	2752	TGCGGAGCCCTGTCCTTCCTACGTACCCAGCCTTGAGAGACTTACTCTTGATTGAACGA	2811
QY	3625	GGATTGTGGAACCTTCTGGAGCGAGGGGTGGGAAGCCCTCAAAATATGTGGGAATTC	3684
Db	2812	GGATTGTGGAACCTTCTGGAGCGAGGGGTGGGAAGCCCTCAAAATATGTGGGAATTC	2871
QY	3685	TACAGTATTGTAGACTCAGAGACTTAAAGATAGTCTGTTAACTTGTCTCATGCGACAGCA	3744
Db	2872	TACAGTATTGTAGACTCAGAGACTTAAAGATAGTCTGTTAACTTGTCTCATGCGACAGCA	2931
QY	3745	TACAGTATTGTAGAGGGACAGATAGGCTTATAGAGATTTACAGCAGCTTATAGAGTA	3804
Db	2932	TACAGTATTGTAGAGGGACAGATAGGCTTATAGAGATTTACAGCAGCTTATAGAGTA	2991
QY	3805	TTGCGCCATACCTAGAGAGATTAAGACAGGGCTTGGAAGAGATTTGCTATAGATGGGT	3864
Db	2992	TTGCGCCATACCTAGAGAGATTAAGACAGGGCTTGGAAGAGATTTGCTATAGATGGGT	3051
QY	3865	GGCAAGTGTCTCAAAAAGTAGTGTGATTGGATGGCTGTAAAGGAAAGATGAGACGA	3924
Db	3052	GGCAAGTGTCTCAAAAAGTAGTGTGATTGGATGGCTGTAAAGGAAAGATGAGACGA	3111
QY	3925	GCTGAGCCACAGCAGATGGGGTGGGAGCAGTATCTCGATCTAGA	3971
Db	3112	GCTGAGCCACAGCAGATGGGGTGGGAGCAGTATCTCGATCTAGA	3158
RESULT 14			
US-10-003-035-21			
; Sequence 21, Application US/10003035			
; Publication No. US20020155127A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Dabner			
; TITLE OR INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS			
; FILE REFERENCE: 22488-712			
; CURRENT APPLICATION NUMBER: US/10/003,035			
; CURRENT FILING DATE: 2001-11-01			
; PRIOR APPLICATION NUMBER: 09/585,599			
; PRIOR FILING DATE: 2000-06-02			
; PRIOR APPLICATION NUMBER: PCT/US01/18238			
; PRIOR FILING DATE: 2001-06-04			
; NUMBER OF SEQ ID NOS: 75			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 21			
; LENGTH: 3417			
; TYPE: DNA			
; ORGANISM: Artificial sequence			
; FEATURE:			
; OTHER INFORMATION: Modified Env/Tat/Rev/Nef from strain BH10			
US-10-003-035-21			
Query Match			
Best Local Similarity 41.3%; Score 2571; DB 13; Length 3417;			
Matches 2630; Conservative 0; Mismatches 45; Indels 12; Gaps 2;			
QY	1291	CCATGAGAGGAGGAGAAAGTATCACCACCTTGTGGAGATGGGGGTGGAAATGGGGCACCA	1350
Db	478	CAATGAGAGGAGGAGAAATATCACACCTTGTGGAGATGGGGGTGGAGATGGGGCACCA	537
QY	1351	TGTCCTTGGGATTTGATGATCTGATGCTAGACAGAAAAATTTGGGTGCACCCGCTATT	1410
Db	538	TGTCCTTGGGATTTGATGATCTGATGCTAGACAGAAAAATTTGGGTGCACAGCTATT	597
QY	1411	ATGGGGACTCTGTGTGAAGGAGAACCAACCACTCTATTTTGTGCATCAGATGCTAAG	1470



OY	3625	GGATTGAGAACTCTGGGACGACAGGGGGTGGGAAGCCCTCAATATATGCGAATCTCC	3668
Dp	2812	GGATTGGAACCTCTGGGAGCGCAGGGGGTGGGAAGCCCTCAATATATGCGAATCTCC	2871
OY	3685	TACAGTATTGGAGTCAGGAACATAAGCTATGCTGTATTACTTCTCAATCCACAGCCA	3744
Dp	2872	TACAGTATTGGAGTCAGGACGCTAAAGAATATGCTGTAGCTTCTCAATCCACAGCTA	2931
OY	3745	TACAGTAGACGACGAGGACAGATAGGGTTATAGAAGTATTCACAGACGCTATAGAGCTA	3804
Dp	2932	TACAGTAGACTCGAAGGGACAGATAGGGTTATAGAAGTAGTACAGAGACCTATAGAGCTA	2992
OY	3805	TTGCGCCACATACCTAGAAAGANTAAAGACAGGGCTTTGGAAAAGATTTTGCATTAAGATGGGT	3864
Dp	2992	TTGCCCCACATACCTAGAAAGAAATAAGACAGGGCTTTGGAAAAGATTTTGCATTAAGATGGGT	3051
OY	3865	GGCAAGTGGTCAAAAATAGTGTGATTGGATGGGCTGCTGTAAAGGAAAGAAATGAGACGA	3924
Dp	3052	GGCAAGTGGTCAAAAATAGTGTGTTGGATGGGCTGCTGTAAAGGAAAGAAATGAGACGA	3111
OY	3925	GCTAGGCCACAGACAGATGGGGTGGGAGCAGTATCTGCAGATCTTACA	3971
Dp	3112	GCTAGGCCACAGACAGATGGGGTGGGAGCAGCATTCTGCAGACCTTACA	3158

```

RESULT 15
US-10-286-332A-24
; Sequence 24, Application US/10286332A
; Publication No. US20030138459A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: METHOD OF VACCINATION THROUGH SEROTYPE ROTATION
; FILE REFERENCE: 22488-738
; CURRENT APPLICATION NUMBER: US/10/286,332A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCP/US01/18238
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 10/003,035
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24
; LENGTH: 2583
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env
; US-10-286-332A-24

```

Query Match	40.0%	Score 2493.6	DB 12	Length 2583
Best Local Similarity	98.3%	Pred. No. 0		
Matches 2533; Conservative	0	Mismatches 39	Indels 6	Gaps 1

OY	1286	CGCCACCATGAGTAGTGAAGGAGAAGTAATCAGCACTTTGTGGAGATGGGGGTGAAATGGGG	1345
Db	6	CGCCACCATGGGATGGAAGGAGAAATATCAGCACCTTTGTGGAGATGGGGGTGGAATGGGG	65
OY	1346	CACCATGCCTCTTGCGATTATTGATGATCTGAGTCGTACAGAAAAAATTTGGTGGCACCGT	1405
Db	66	CACCATGCTCTTGCGATGTTTGAATGATCTGTAGTGTCTACAGAAAAAATTTGTGGTGCACAGT	125
OY	1406	CTATTATGGGGTACTCTGTGTGAAGAAGAACCAACCACTCTATTATTTGTGCATCAGATGC	1465
Db	126	CTATTATGGGGTACTCTGTGTGAAGAAGAACCAACCACTCTATTATTTGTGCATCAGATGC	185
OY	1466	TAAAGCATATGATACAGAGGTACATATANGTTTTGGGCCAACATCCCGTGTACCCACAGA	1525
Db	186	TAAAGCATATGATACAGAGGTACATATANGTTTTGGGCCAACATCCCGTGTACCCACAGA	245
OY	1526	CCCCACCCCACAGAAAGTAGTATTGTGTAATGTGTACAGAAAATTTTAACATGTGTGAAAAA	1585

Db	246	CCCCACCCACAGAGAGTAGTATTGGTAAAGTGCACGAAAAATTTTACACATGTGAAAAA	305
QY	1586	TGACATGGTAAACAGATGCATGAGAGATATATACGTTTATGGATCAAAAGCTTAAAGCC	1645
Db	306	TGACATGGTAGACAGATGCATGAGAGATATATACGTTTATGGATCAAAAGCTTAAAGCC	365
QY	1646	ATGTGTAAATTAACCCCACTGTGTGTTAGTTTAAAGTGCACGTATTTGAAAGATGATAC	1705
Db	366	ATGTGTAAATTAACCCCACTGTGTGTTAGTTTAAAGTGCACGTATTTGAAAGATGATAC	425
QY	1706	TAATACCAATAGTAGTAGCGGAGACATGATATGAGAGAAAGAGATAAAAAACTGCTC	1765
Db	426	TAATACCAATAGTAGTAGCGGAGACATGATATGAGAGAAAGAGATAAAAAACTGCTC	485
QY	1766	TTTCAATATCACGACACACATPAGAGATAGGTGCAGAAAGATATGCAATTCCTTTATPA	1825
Db	486	TTTCAATATCACGACACACATPAGAGATAGGTGCAGAAAGATATGCAATTCCTTTATPA	545
QY	1826	ACTTGATATTAGTACCAATAGATAA-----TACCAGATATTAGTTGATAAATTGTAACAC	1879
Db	546	ACTTGATATTAGTACCAATAGATAA-----TACCAGATATTAGTTGATAAATTGTAACAC	605
QY	1880	CTCAGTCATPACACAGGCGCTGCCAAGAGTATCCTTTAGGCCAATTCCTCATCATATTATG	1939
Db	606	CTCAGTCATPACACAGGCGCTGCCAAGAGTATCCTTTAGGCCAATTCCTCATCATATTATG	665
QY	1940	TGCCCCGGCTGGTTTGGGATTCTPAAATATGATATTAAGACGTTCAATGGAACAGGACC	1999
Db	666	TGCCCCGGCTGGTTTGGGATTCTPAAATATGATATTAAGACGTTCAATGGAACAGGACC	725
QY	2000	ATGTACAATGTCCAGCACAGTACACATGTACACATGGAATTCAGGCGCAGTAGATCAACCA	2058
Db	726	ATGTACAATGTCCAGCACAGTACACATGTACACATGGAATTTAGGCCAGTAGATCAACCA	785
QY	2060	ACTGCTGTTAAATGGCAGCTPACGAGAAAGATGTAGTAAATGTAGATCTGCCAATTTAC	2119
Db	786	ACTGCTGTTAAATGGCAGCTPACGAGAAAGATGTAGTAAATGTAGATCTGCCAATTTAC	845
QY	2120	AGACAATGCTPAAACCATATATGTACACTGACACATCTGTAGAAATTAATTTGTACAG	2179
Db	846	AGACAATGCTPAAACCATATATGTACACTGACACATCTGTAGAAATTAATTTGTACAG	905
QY	2180	ACCACAACAATACAGAAAAAGTATCCGATCCAGAGGGGACCCAGGAGACATTTGT	2239
Db	906	ACCACAACAATACAGAAAAAGTATCCGATCCAGAGGGGACCCAGGAGACATTTGT	965
QY	2240	TACCAATAGGAAAAATAGCAATATGAGACAAAGCACATTTGTAACTATAGACAAATG	2299
Db	966	TACCAATAGGAAAAATAGCAATATGAGACAAAGCACATTTGTAACTATAGACAAATG	1025
QY	2300	GAATGCCACTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATTAATTAAC	2359
Db	1026	GAATGCCACTTAAACAGATAGATAGCAAAATTAAGAGAACAAATTTGGAATTAATTAAC	1085
QY	2360	AATAATCTTTAAGCAATCCTCAGAGAGGGACCCAGAAATGTAAACGACAGTTTAAATG	2419
Db	1086	AATAATCTTTAAGCAATCCTCAGAGAGGGACCCAGAAATGTAAACGACAGTTTAAATG	1145
QY	2420	TGGAGGGGAATTTTCTACGTGAATTAATCAACCACTGTTTATATGTACTTGGTTTAAAG	2479
Db	1146	TGGAGGGGAATTTTCTACGTGAATTAATCAACCACTGTTTATATGTACTTGGTTTAAAG	1205
QY	2480	TACTTGGAGTACTGAAGGGTCAAAATACACGTGAAGGAATGACACAATCACACTCCATG	2539
Db	1206	TACTTGGAGTACTGAAGGGTCAAAATACACGTGAAGGAATGACACAATCACACTCCATG	1265
QY	2540	CAGAAATAAACAATTTAATAACATGTGGCAGGAAGTGAAGAAACAATGTATGCCCTTCC	2599
Db	1266	CAGAAATAAACAATTTAATAACATGTGGCAGGAAGTGAAGAAACAATGTATGCCCTTCC	1325
QY	2600	CATCAGTGACAAATAGATGTTTCATCAAAATTAATCTAGCGGTGCTATTAAACAGATAGG	2659

Db	1326	CATCAGTGGGCAAAATTATGATGTTTCATCAATATATTTACAGGGCTGCTATTACACAGAGATGG	1385
Qy	2660	TGTAATATACACAAATGGGTCCGAGATCTTCACAGCTCGAGAGAGGCGATATGAGGACAA	2719
Db	1386	TGGTAATATGCAACAAATGATCCGAGATCTTCAGACCTCGAGAGAGGAGATATGAGGACAA	1445
Qy	2720	TTGGAGAAAGGAAATTATATTAATATTAAGTAGTAATAATTGAACCTTTAGAGATGACAC	2779
Db	1446	TTGGAGAAAGGAAATTATATTAATTAAGTAGTAATAATTGAACCTTTAGAGATGACAC	1505
Qy	2780	CACCAAGGCAAAAGAGAGATGCTGACAGAGAAAAAAGAGAGTGGGAATGAGAGCTTT	2839
Db	1506	CACCAAGGCAAAAGAGAGATGCTGACAGAGAAAAAAGAGAGTGGGAATGAGAGCTTT	1565
Qy	2840	GTTCTCTTGGGTCTTGGGAGCAGACAGAAAGCACTATGGGCTGACGTCATGACGCTGAC	2899
Db	1566	GTTCTCTTGGGTCTTGGGAGCAGACAGAAAGCACTATGGGCTGACGTCATGACGCTGAC	1625
Qy	2900	GGTACAGGCCAGACAAATATATGCTGATATAGTGCAGCAGCAACAAATTTGCTGAGGGC	2959
Db	1626	GGTACAGGCCAGACAAATATATGCTGATATAGTGCAGCAGCAACAAATTTGCTGAGGGC	1685
Qy	2960	TATGAGGCCCAACAGACATCTGTCACACAGCTGGGGCATCAACAGCTCCAGGC	3019
Db	1686	TATGAGGCCCAACAGACATCTGTCACACAGCTGGGGCATCAACAGCTCCAGGC	1745
Qy	3020	AAGAATCTGGCTGTGGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTG	3079
Db	1746	AAGAATCTGGCTGTGGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTG	1805
Qy	3080	CTCTGGAAAACTATTTGGCACCCACTGCTGGCTTGGAAATGCTAGTTGGAGTAATAATC	3139
Db	1806	CTCTGGAAAACTATTTGGCACCCACTGCTGGCTTGGAAATGCTAGTTGGAGTAATAATC	1865
Qy	3140	TCTGGAACAGATTTGGAAATATACATGACCTGGATGGAGTGGGACAGAAATTAACAATTA	3199
Db	1866	TCTGGAACAGATTTGGAAATATACATGACCTGGATGGAGTGGGACAGAAATTAACAATTA	1925
Qy	3200	CACAAGCTTAATACACTCTTAATTGAAGATCGCAAAACCAAGCAAAAGAAATGAACA	3259
Db	1926	CACAAGCTTAATACACTCTTAATTGAAGATCGCAAAACCAAGCAAAAGAAATGAACA	1985
Qy	3260	AGAATTTTGGAAATAGATTAATATGGCAAGTTTGGAAATTTGTTTACATTAACAATTTG	3319
Db	1986	AGAATTTTGGAAATAGATTAATATGGCAAGTTTGGAAATTTGTTTACATTAACAATTTG	2045
Qy	3320	GCTGTGATATATAAATTTATTAATGATAGAGAGGCTTGGTATGATTAAGATAGT	3379
Db	2046	GCTGTGATATATAAATTTATTAATGATAGAGAGGCTTGGTATGATTAAGATAGT	2105
Qy	3380	TTTTGCTGTACTTCTAATAGTAATAGATTAGGCAAGGATATTCACCATTAATCTTTCA	3439
Db	2106	TTTTGCTGTACTTCTAATAGTAATAGATTAGGCAAGGATATTCACCATTAATCTTTCA	2165
Qy	3440	GACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGATGAAAGAAAGTTGG	3499
Db	2166	GACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGATGAAAGAAAGTTGG	2225
Qy	3500	AGAGAGAGACAGAGACGATCCATTGATAGTAAGAGATTCCTTAGCATATCTGGGA	3559
Db	2226	AGAGAGAGACAGAGACGATCCATTGATAGTAAGAGATTCCTTAGCATATCTGGGA	2285
Qy	3560	CGATCTCGGAGGCTGCTGCTTCACGCTACACCGCTGAGAGACTTACTCTGATTTGT	3619
Db	2286	CGATCTCGGAGGCTGCTGCTTCACGCTACACCGCTGAGAGACTTACTCTGATTTGT	2345
Qy	3620	AACGAGATTGTGGAACCTTCTGGGACGACAGGGGTGGAAAGCCCTCAATATTTGGTGA	3679
Db	2346	AACGAGATTGTGGAACCTTCTGGGACGACAGGGGTGGAAAGCCCTCAATATTTGGTGA	2405
Qy	3680	TCTTCTACAGTATTTGGAGTACGAACTAAAGAAATAGTCTGTTAACTTGTCTAATGCCAC	3739
Db	2406	TCTTCTACAGTATTTGGAGTACGAACTAAAGAAATAGTCTGTTAACTTGTCTAATGCCAC	2465

Qy

3740

AGCCATAGAGTAGTCTGAGGGACAGATAGGGTTATAGAGTATTTACAGCAGCTTATAG

3799

Db

2466

AGCTATAGAGTAGTCTGAGGGGACAGATAGGGTTATAGAGTATTTACAGCAGCTTATAG

2525

Qy

3800

AGCTATTTCCCAATACCTAGAAAGAAATTAAGACAGGCTTGGAAAGGATTTTGGCTATTA

3857

Db

2526

AGCTATTTCCCAATACCTAGAAAGAAATTAAGACAGGCTTGGAAAGGATTTTGGCTATTA

2583

Search completed: September 17, 2003, 22:52:22

Job time : 1376 secs

Search completed: September 17, 2003, 22:52:22  
Job time : 1376 secs

:52:22

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:35:16 : Search time 11656 Seconds  
(without alignments)  
12988.382 Million cell updates/sec

Title: US-09-913-159A-10  
Perfect score: 6229  
Sequence: 1 ctgacgcgcctctgacgcgc.....attccccaagaatgcac 6229

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1055.8	16.9	1070	9	AJ281552 4A3A-P6F1
2	986.4	15.8	1013	12	BM438846 IPLVR0015
3	913.4	14.7	973	14	CD458281 Fg08_08h0
4	902	14.5	1004	9	AJ281480 4A3A-P4G8

Result No.	Score	Query Match	Length	ID	Description
5	882.8	14.2	917	14	CD458286 Fg08_09a0
6	882.8	14.2	956	12	AL959357
7	870.2	14.0	935	12	BG838279
8	862.6	13.8	918	14	CD459092 Fg08_08e0
9	851.4	13.7	986	11	CNS08PE5
10	841	13.5	841	9	AL042026
11	834.4	13.4	915	11	CNS09072
12	832.8	13.4	872	14	CD450885
13	824.6	13.2	854	12	CD450885
14	822.4	13.2	893	14	BM438950
15	816.8	13.1	970	9	AL879792
16	814.4	13.0	1126	29	CD458333 Fg08_09e0
17	807	12.8	863	11	CNS09LFC
18	797.6	12.8	1011	29	B2577702
19	774.8	12.4	1336	29	B2575810
20	774.4	12.4	789	14	CD280920
21	774.4	12.4	1574	29	B2572566
22	767.6	12.3	1370	29	B2571721
23	763.2	12.3	780	13	BQ825693
24	757	12.2	759	14	CD279661
25	755.4	12.1	1463	29	B2571475
26	747.8	12.0	800	9	AJ281449
27	747.2	12.0	802	11	CNS08WC8
28	743.6	11.9	832	12	BG923768
29	742.6	11.9	786	14	CD458721
30	741	11.9	741	14	CD279174
31	738.2	11.9	966	29	B2575002
32	737.8	11.8	1003	29	B2576686
33	736.8	11.8	950	29	B2571129
34	733.8	11.8	872	11	CNS093V9
35	733.4	11.8	817	11	CNS09DS1
36	730.2	11.7	998	29	B2576702
37	730	11.7	730	14	CD281097
38	729.2	11.7	840	29	B2571995
39	729	11.7	729	14	CD279546
40	728.6	11.7	842	12	B1687610
41	728	11.7	728	14	CD279322
42	726.2	11.7	881	29	B2577222
43	726	11.7	726	14	CD281811
44	722.6	11.6	765	11	CNS093CH
45	722	11.6	1394	29	B2580042

## ALIGNMENTS

RESULT 1  
AJ281552  
LOCUS 1070 bp mRNA linear EST 30-JUN-2000  
DEFINITION 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A-P6F11, mRNA sequence.

ACCESSION AJ281552  
VERSION AJ281552.1 GI:6929432

KEYWORDS EST.  
SOURCE Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

REFERENCE 1 (bases 1 to 1070)  
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M. B. and Kafatos, F. C.

ANOPHELES GAMBIAE PILOT GENE DISCOVERY PROJECT: IDENTIFICATION OF MOSQUITO INNATE IMMUNITY GENES FROM EXPRESSED SEQUENCE TAGS generated from immune-competent cell lines

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

COMMENT 10841361  
Contact: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerothofstrasse 1, 69117 Heidelberg, Germany.

FEATURES  
Source

Location/Qualifiers  
1. .1070  
/organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P6F11"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoRI; Site\_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1986) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 263 a 283 c 255 g 269 t

ORIGIN

Query Match 16.9% Score 1055.8; DB 9; Length 1070;  
Best Local Similarity 99.7%; Pred. No. 5.1e-199;  
Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

4471 TTGCTGGCGTTTTCATAGAGCTCGCCGCCCTGACGACATCAAAAAATCGACGCTCA 4530  
1 TTGCTGGCGTTTTCATAGAGCTCGCCGCCCTGACGACATCAAAAAATCGACGCTCA 60  
4531 AGTCAGAGGTGGCCAAACCCGACAGACTATTAAGATACAGCGGTTTCCCTCGGAAGC 4550  
61 AGTCAGAGGTGGCCAAACCCGACAGACTATTAAGATACAGCGGTTTCCCTCGGAAGC 120  
4591 TCCCTCGGCGCTCCCTGTTCCACCCCTGCGCTTACCGGATACGTGCTCCGCTTCTC 4650  
121 TCCCTCGGCGCTCCCTGTTCCACCCCTGCGCTTACCGGATACGTGCTCCGCTTCTC 180  
4651 CTTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGATCTCAGTTCGGGTAG 4710  
181 CTTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGATCTCAGTTCGGGTAG 240  
4711 GTCTGTGGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTCAGCCGCGCTGGCC 4770  
241 GTCTGTGGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTCAGCCGCGCTGGCC 300  
4771 TTATCCGGTAACATCTGCTTGTAGTCCAAACCGGTAAAGACAGCTTATCGCCATGGCA 4830  
301 TTATCCGGTAACATCTGCTTGTAGTCCAAACCGGTAAAGACAGCTTATCGCCATGGCA 360  
4831 GCAGCCACTGTAAACAGGATTAGCAGAGCGAGTATGAGCGGTGTACAGAGTTCTTG 4890  
361 GCAGCCACTGTAAACAGGATTAGCAGAGCGAGTATGAGCGGTGTACAGAGTTCTTG 420  
4891 AAGTGGGGCTACTACGCTAGCTAGCAAGAGACAGTATTTGATCTGGCTGTCTG 4950  
421 AAGTGGGGCTACTACGCTAGCTAGCAAGAGACAGTATTTGATCTGGCTGTCTG 480  
4951 AAGCAGTACCTTCGAAAAAGAGTGTAGTCTTGTATCCGCAAAACAACACCGCT 5010  
481 AAGCAGTACCTTCGAAAAAGAGTGTAGTCTTGTATCCGCAAAACAACACCGCT 540  
5011 GGTAGCGGTGTTTTTTTGTTCGACGACAGATTACGCGCAGAAAAAAGATCTCAA 5070  
541 GGTAGCGGTGTTTTTTTGTTCGACGACAGATTACGCGCAGAAAAAAGATCTCAA 600  
5071 GAAATGCTTTGATCTTTTACGGGCTTGACGCTAGTGAAGCAAAACTCAGCTTAA 5130  
601 GAAATGCTTTGATCTTTTACGGGCTTGACGCTAGTGAAGCAAAACTCAGCTTAA 660  
5131 GGAATTTTGGTCATGAGATTATCAAAAAGATCTTCACCTAGATCTTTTAATTAAAA 5190  
661 GGAATTTTGGTCATGAGATTATCAAAAAGATCTTCACCTAGATCTTTTAATTAAAA 720

QY 5191 TGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACCTGCTGACAGTTACCAATGC 5250  
DB 721 TGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACCTGCTGACAGTTACCAATGC 780  
QY 5251 TTATCACTGAGGACCTATCTCAGCATCTGTCTATTTTGGTATCCATAGTCCCTGA 5310  
DB 781 TTATCACTGAGGACCTATCTCAGCATCTGTCTATTTTGGTATCCATAGTCCCTGA 840  
QY 5311 CTCCCGCTGCTGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 5370  
DB 841 CTCCCGCTGCTGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 900  
QY 5371 ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATTCAGCAATTAACACGACGCC 5430  
DB 901 ATGATACCGGAGACCCACGCTCACCGCTCCAGATTATTCAGCAATTAACACGACGCC 960  
QY 5431 GGAAGGCCGAGCCAGCAAGTGTCTGCAACTTATTCGCTTCATTCACATCTATTAT 5490  
DB 961 GGAAGGCCGAGCCAGCAAGTGTCTGCAACTTATTCGCTTCATTCACATCTATTAT 1019  
QY 5491 TGTGCGCGGAGCTAGATAGTATGTTCCGCTTATATGTTGGCAGC 5541  
DB 1020 TGTGCGCGGAGCTAGATAGTATGTTCCGCTTATATGTTGGCAGC 1070

RESULT 2  
BM438846/c 1013 bp mRNA linear EST 31-JAN-2002  
LOCUS IPIV00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA  
DEFINITION sequence.  
ACCESSION BM438846  
VERSION BM438846.1 GI:18460568  
KEYWORDS EST.  
SOURCE Ictalurus punctatus (channel catfish)  
ORGANISM Ictalurus punctatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes; Ictaluridae; Ictalurus.  
REFERENCE 1 (bases 1 to 1013)  
Feng J., Kucuktas H., Kocbas A., Li P. and Liu Z.  
Transcriptome of channel catfish (Ictalurus punctatus): Initial analysis of expressed sequence tags from the liver  
JOURNAL Unpublished  
COMMENT Contact: Liu ZJ  
The Fish Molecular Genetics and Biotechnology Laboratory,  
Department of Fisheries and Allied Aquacultures and Program of Cell  
and Molecular Biosciences  
Auburn University  
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
Tel: 334 844 4054  
Fax: 334 844 9208  
Email: zliu@cesag.auburn.edu  
Seq primer: M13 Reverse.  
FEATURES  
Source  
1. .1013  
/organism="Ictalurus punctatus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7998"  
/clone\_lib="Liver cDNA library"  
/note="Organ: Liver; Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 273 a 228 c 245 g 266 t 1 others

ORIGIN

Query Match 15.8%; Score 986.4; DB 12; Length 1013;  
Best Local Similarity 99.5%; Pred. No. 3.1e-185;  
Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 5044 ATTACGGCAGAAAAAAGATCTCAAGAGATCTTGTATCTTTTACGGGCTGAC 5103  
DB 1013 ATTACGGCAGAAAAAAGATCTCAAGAGATCTTGTATCTTTTACGGGCTGAC 954  
QY 5104 GCTCAGTGAACGAACCTCAGCTTAAGGAGTTTGGTCATGAGATTATCAAAAAGATC 5163



```

Db      953 GCTAGTGGAAAGAACTCAAGTTAAGGATTTTGGTCATGAGATTTCAAAAAGATC 894
Qy      5164 TTCACGTGATCCTTTTAATTAATAAAGATTAAATCAATCTAAAGTATATAG 5223
Db      893 TTCACGTGATCCTTTTAATTAATAAAGATTAAATCAATCTAAAGTATATAG 834
Qy      5224 TAACTTGGTGCAGTTACCAATGCTTATTCAGTGGGACCTATCTCAGGATCTGT 5283
Db      833 TAACTTGGTGCAGTTACCAATGCTTATTCAGTGGGACCTATCTCAGGATCTGT 774
Qy      5284 CTATTCCTTCATCATAGTTAGTCCCTGACTCCCGCTGCTGATAGTACTAGTACGAGAG 5343
Db      773 CTATTCCTTCATCATAGTTAGTCCCGTACTCCCGCTGCTGATAGTACTAGTACGAGAG 714
Qy      5344 GCGTTACCATCTGGCCCCAGTGCAGTCAATGATPACCGGAGAACCCAGCTCACCGGCTCCA 5403
Db      713 GCGTTACCATCTGGCCCCAGTGCAGTCAATGATPACCGGAGAACCCAGCTCACCGGCTCCA 654
Qy      5404 GATTATACGCAATTAACGACGACGCGGAAGGCGGACGACGAGAGTGTCTGCAACT 5463
Db      653 GATTATACGCAATTAACGACGACGCGGAAGGCGGACGACGAGAGTGTCTGCAACT 594
Qy      5464 TTATCCGCGCTCCATCCAGTCTATTATATGTTGCCGGAAGTAGATAGTATGTTCCGCA 5523
Db      593 TTATCCGCGCTCCATCCAGTCTATTATATGTTGCCGGAAGTAGATAGTATGTTCCGCA 534
Qy      5524 GTTAAATAGTTGGCCGACGTTGTTGCCATTCTCTACAGGACATCTGCTGCTGCTGCTG 5583
Db      533 GTTAAATAGTTGGCCGACGTTGTTGCCATTCTCTACAGGACATCTGCTGCTGCTGCTG 474
Qy      5584 TTTGGTATGGCTTCATTCAGTCCGCTGCCAGAGATCAAGGCGAGTTACA-TGATCCCC 5642
Db      473 TTTGGTATGGCTTCATTCAGTCCGCTGCCAGAGATCAAGGCGAGTTACA-TGATCCCC 414
Qy      5643 CATGTTGTGCAAAAAAGCGGTTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5702
Db      413 CATGTTGTGCAAAAAAGCGGTTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355
Qy      5703 GCGCCGACGTTATTCATCATAGTGTATGACAGACATGATATCTCTTACTGTCTGCTGCTG 5762
Db      354 GCGCCGACGTTATTCATCATAGTGTATGACAGACATGATATCTCTTACTGTCTGCTGCTG 295
Qy      5763 ATCCGTAGATGCTTTCTGAGCTGCTGAGTCAACCAAGTATCTGAGATTAAGT 5822
Db      294 ATCCGTAGATGCTTTCTGAGCTGCTGAGTCAACCAAGTATCTGAGATTAAGT 235
Qy      5823 TATGCGGACGAGATGCTCTGCGCGGCTCAATACGAGTAAATACCGGCGCCATAG 5882
Db      234 TATGCGGACGAGATGCTCTGCGCGGCTCAATACGAGTAAATACCGGCGCCATAG 175
Qy      5883 CAGAACTTTAAAAAGTGTCTCATCTATTGAAAAAGTCTTCTGCGGCGGAAAACTCTCAAGAT 5942
Db      174 CAGAACTTTAAAAAGTGTCTCATCTATTGAAAAAGTCTTCTGCGGCGGAAAACTCTCAAGAT 115
Qy      5943 CTTAACCGCTGTTGAGATCCATGTTGATTAACCCACTGTGTCAACCACTGATCTTACG 6002
Db      114 CTTAACCGCTGTTGAGATCCATGTTGATTAACCCACTGTGTCAACCACTGATCTTACG 55
Qy      6003 ATCTTTACTTTTCAACGAGCTTTCTGCGGTGAGCAAAAAAGAGAGCAAAATGC 6056
Db      54 ATCTTTACTTTTCAACGAGCTTTCTGCGGTGAGCAAAAAAGAGAGCAAAATGC 1

```

```

RESULT 3
CD458281/ 973 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08h09_R Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08h09, mRNA sequence.
ACCESSION CD458281
VERSION CD458281.1 GI:31373021
KEYWORDS EST.
SOURCE Gibberella zeae

```

```

ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 973)
AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori
,J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Spott,D. and
Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL Unpublished
COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.
FEATURES
location/Qualifiers
source 1..973
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="E:taxon:5518"
/clone="Fg08_08h09"
/tissue_type="mycelium"
/dev_stage="asexual"
/lab_host="E. coli DH10B"
/clone_11b="Fg08_AAFC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/note="Vector: pBluescript II+; Site.1: EcoRI; Site.2:
XhoI; Fusarium graminearum grown on a complex plant
substrate- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."
BASE COUNT 264 a 215 c 238 g 251 t
ORIGIN
Query Match 14.7% Score 913.4; DB 14; Length 973;
Best Local Similarity 98.7%; Pred. No. 9,7e-171;
Matches 925; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 5293 TCATCATAGTTGCTGCTGACTCCCGCTGATAGTAACTAGCATGAGGAGGCTTACCA 5352
Db 965 TCATCATAGTTGCTGCTGACTCCCGCTGATAGTAACTAGCATGAGGAGGCTTACCA 907
Qy 5353 TCTGCCCCAGTGTGCAATGATATACCGGAGACCCAGCTCACCGCTCCAGATTATCA 5412
Db 906 TCTGCCCCAGTGTGCAATGATATACCGGAGACCCAGCTCACCGCTCCAGATTATCA 847
Qy 5413 GCAATTAACCGACCGCGGAGGCGGAGGCGGAGAGTGTCTGCAACTTTATCGCC 5472
Db 846 GCAATTAACCGACCGCGGAGGCGGAGGCGGAGAGTGTCTGCAACTTTATCGCC 787
Qy 5473 TCCATCCAGTCTATTAATTGTTGGCGGAGCTAGTAAGTATGCCAGTTATAGT 5532
Db 786 TCCATCCAGTCTATTAATTGTTGGCGGAGCTAGTAAGTATGCCAGTTATAGT 727
Qy 5533 TTGCGCAAGTGTGTTGCCATTGCTACAGAGCATGCTGTGTCACGCTCTGCTTGGTATG 5592
Db 726 TTGCGCAAGTGTGTTGCCATTGCTACAGAGCATGCTGTGTCACGCTCTGCTTGGTATG 667
Qy 5593 GCTTATTCACGCTCGGTTCCCAAGATCAAGGAGTATCATATATCCCATGTTGTGC 5652
Db 666 GCTTATTCACGCTCGGTTCCCAAGATCAAGGAGTATCATATATCCCATGTTGTGC 607
Qy 5653 AAAAAGCGGTAGTCTCTGCTGCTCCGATCGTTGACAGAGTAAGTGGCGGAG 5712
Db 606 AAAAAGCGGTAGTCTCTGCTGCTCCGATCGTTGACAGAGTAAGTGGCGGAG 547
Qy 5713 TTATCACTCATGTTATGAGACACTGATATATCTTACTGTCATCCATCCGTAGA 5772
Db 546 TTATCACTCATGTTATGAGACACTGATATATCTTACTGTCATCCATCCGTAGA 487

```

5773 TGCCTTCTGTGACTGTGAGTACTCAACCAAGTCATTGTGAGAAATGATGAGCGGCA 5832  
|||||  
486 TGCCTTCTGTGACTGTGAGTACTCAACCAAGTCATTGTGAGAAATGATGAGCGGCA 427  
5833 CCGAGTTCCTTGTGCGCGCGCTCAATACGAGTATATACCGGGCCACATAGCAAGACTTA 5892  
|||||  
426 CCGAGTTCCTTGTGCGCGCGCTCAATACGAGTATATACCGGGCCACATAGCAAGACTTA 367  
5893 AAAATGCTCATTCATTGAGAAACGTTCTCGGGGCAAAACCTCAAGAGTTTACCGCTG 5952  
|||||  
366 AAAGTCTCATTCATTGAGAAACGTTCTCGGGGCAAAACCTCAAGAGTTTACCGCTG 307  
5953 TTGAGATCCAGTTGATGATTAACCACTGTCACCAAGTATCTTCAAGATCTTTACT 6012  
|||||  
306 TTGAGATCCAGTTGATGATTAACCACTGTCACCAAGTATCTTCAAGATCTTTACT 247  
6013 TTACACAGCTTCTGAGTGAAGCAAAACAGAAAGGCAAAATGCGCAAAAGGAGTA 6072  
|||||  
246 TTACACAGCTTCTGAGTGAAGCAAAACAGAAAGGCAAAATGCGCAAAAGGAGTA 187  
6073 AGGCGACACGAAATGTTGAATCTCATCTCTCTTTTCAATATTTGAAGCAT 6132  
|||||  
186 AGGCGACACGAAATGTTGAATCTCATCTCTCTTTTCAATATTTGAAGCAT 127  
6133 TATCAGGTTTATGTTCTCATAGCGGATACATATTTGAATATTTAAGAAATTAACAA 6192  
|||||  
126 TATCAGGTTTATGTTCTCATAGCGGATACATATTTGAATATTTAAGAAATTAACAA 67  
6193 ATAGGGGTTCCGCGACATTTCCCGCAAAAGTGCCAC 6229  
|||||  
66 ATAGGGGTTCCGCGACATTTCCCGCAAAAGTGCCAC 30

RESULT 4  
AJ281480 1004 bp mRNA linear EST 30-JUN-2000  
LOCUS  
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles  
gambiae cDNA clone 4A3A-P4G8, mRNA sequence.  
ACCESSION  
AJ281480  
VERSION  
AJ281480.1 GI:6929360  
KEYWORDS  
EST.  
SOURCE  
Anopheles gambiae (African malaria mosquito)  
ORGANISM  
Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Anopheles.  
REFERENCE  
1 (bases 1 to 1004)  
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,  
Dorohoe, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M. B.  
and Kafatos, F. C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
location/qualifiers  
1. 1004  
/organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P4G8"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/clone\_lib="Anopheles gambiae immune competent 4A3A"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: EcoRI, Site\_2: NotI; sequenced from  
forward priming site which reads from the 3' end of the

CDNA. The 4A3A is a directionally cloned and normalized  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares  
(1996) : Normalization and Subtraction: Two approaches to  
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others  
ORIGIN

Query Match 14.5%; Score 902; DB 9; Length 1004;  
Best Local Similarity 99.0%; Pred. No. 1.8e-168;  
Matches 979; Conservative 1; Mismatches 2; Indels 7; Gaps 7;

4464 GCGCGCTGTGCTGCGCTTTTCCATPAGGCTCGCCCTGAGCAGATCCAAAATCG 4523  
|||||  
23 GCGCGCTGTGCTGCGCTTTTCCATPAGGCTCGCCCTGAGCAGATCCAAAATCG 82  
4524 ACGCTCAAGTACAGAGTGGGCAACCCGACAGGACTTAAAGATACAGAGGTTCCCGC 4583  
|||||  
83 ACGCTCAAGTACAGAGTGGGCAACCCGACAGGACTTAAAGATACAGAGGTTCCCGC 142  
4584 TGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATCTGTCCGC 4643  
|||||  
143 TGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATCTGTCCGC 202  
4644 CTTTCTCCCTTCCGGAACCGTGGCGCTTTCATAGCTCAGCGTGTAGTCTCACTTC 4703  
|||||  
203 CTTTCTCCCTTCCGGAACCGTGGCGCTTTCATAGCTCAGCGTGTAGTCTCACTTC 262  
4704 GGTGTAGTGTGTTGCTCCCAAGCTGGGCTGTGTCCAGAACCCCGCTTACGCCGACG 4763  
|||||  
263 GGTGTAGTGTGTTGCTCCCAAGCTGGGCTGTGTCCAGAACCCCGCTTACGCCGACG 322  
4764 CTGGCCCTTATCCGTAACATATCTTGAATCCCAACCCGTAAGACAGCATTTATGCC 4823  
|||||  
323 CTGGCCCTTATCCGTAACATATCTTGAATCCCAACCCGTAAGACAGCATTTATGCC 382  
4824 ACTGGACAGCCACTGTGTAACAGATTTAGACAGCGAGTATAGCGGTGTCTACAGA 4883  
|||||  
383 ACTGGACAGCCACTGTGTAACAGATTTAGACAGCGAGTATAGCGGTGTCTACAGA 442  
4884 GTTCTTAAGTGGGCGCTTACCTACGCTACATAGAGACAGTATTTGGTATCTCGC 4943  
|||||  
443 GTTCTTAAGTGGGCGCTTACCTACGCTACATAGAGACAGTATTTGGTATCTCGC 502  
4944 TGTCTGAAGCAGTACCTTTCGAAAAAGAGTTGTAAGCTTTGATCCGCAAAAC 5003  
|||||  
503 TGTCTGAAGCAGTACCTTTCGAAAAAGAGTTGTAAGCTTTGATCCGCAAAAC 562  
5004 CACCGCTGTAGCGGTGTTTGTGTAACACACAGATTAACGGGCAAAAAAAG 5063  
|||||  
563 CACCGCTGTAGCGGTGTTTGTGTAACACACAGATTAACGGGCAAAAAAAG 622  
5064 ATCTCAAGAGATCTTGTGATCTTCTAGCGGGTGTAGCGCTGTAGCGGCAAAATC 5123  
|||||  
623 ATCTCAAGAGATCTTGTGATCTTCTAGCGGGTGTAGCGCTGTAGCGGCAAAATC 682  
5124 ACGTTAAGGATTTTGTATGATGATTAACAAAAGATCTTCACTAGATCTTTTAA 5183  
|||||  
683 ACGTTAAGGATTTTGTATGATGATTAACAAAAGATCTTCACTAGATCTTTTAA 742  
5184 TTAATAATGAAGTTTAAATCAATCTAAAGATATATAGATTAACCTGTGTGACACTTA 5243  
|||||  
743 TTAATAATGAAGTTTAAATCAATCTAAAGATATATAGATTAACCTGTGTGACACTTA 802  
5244 CCAATGCTTATACGATGAGGACCATATCTAGCATCTGTCTATTTGTTTCACTCATAGT 5303  
|||||  
803 CCAATGCTTATACGATGAGGACCATATCTAGCATCTGTCTATTTGTTTCACTCATAGT 861  
5304 TGCCTAGTCCCGCTGTGTAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 5363  
|||||  
862 TGCCTAGTCCCGCTGTGTAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 919  
5364 TGCCTAGTATGATACCGGAGACCCAGCTACCGGCTCCAGATTTATACGATTAACCA 5423

Db 920 TGCTCAATGATACCGCGAGA-CAACGGTCA-CGGCTCCAGATT-TTACGACATTAACCA 976  
 QY 5424 GCCACCGGAGAGCGCCGACGCGAGAAGTG 5452  
 Db 977 GCCACCGGAGGAA-GGCCGAGCGCAGAGATG 1004

RESULT 5  
 CD458286/c 917 bp mRNA linear EST 03-JUN-2003  
 LOCUS Fg08\_09a03\_R Fg08\_AAFRC\_ECORC\_Fusarium\_graminearum\_complex\_substrate  
 DEFINITION Gibberella zeae cDNA clone Fg08\_09a03, mRNA sequence.  
 ACCESSION CD458286  
 VERSION CD458286.1 GI:31373026  
 KEYWORDS EST.  
 ORGANISM Gibberella zeae  
 SOURCE Gibberella zeae  
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocryomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 REFERENCE 1 (bases 1 to 917)  
 AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori  
 J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and  
 Tinker,N.A.  
 TITLE A cDNA library prepared from Fusarium graminearum grown on a  
 complex plant substrate  
 JOURNAL Unpublished  
 COMMENT Contact: Watson, Robert.J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1655  
 Fax: (613) 759-1701  
 Email: watsonrj@agr.gc.ca.

FEATURES  
 source  
 1. 917  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg08\_09a03"  
 /tissue\_type="Mycelium"  
 /dev\_stage="Asexual"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="Fg08\_AAFRC\_ECORC\_Fusarium\_graminearum\_complex\_s  
 ubstrate"  
 /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:  
 XhoI; Fusarium graminearum grown on a complex plant  
 substrate--wheat leaves treated to remove most of the low  
 molecular weight, water-soluble components."  
 BASE COUNT 249 a 207 c 220 g 241 t  
 ORIGIN

Query Match 14.2% Score 882.8; DB 14; Length 917;  
 Best Local Similarity 99.8%; Pred. No. 1,le-164;  
 Matches 884; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5344 GCGTTACCATCTGGCCCAAGTGTCTCATGATGATCCGACAGACCCAGCTCAACGGCTCCA 5403  
 Db 917 GCGTTACCATCTGGCCCAAGTGTCTCATGATGATCCGACAGACCCAGCTCAACGGCTCCA 858  
 QY 5404 GATTATATCAGCATTAACACGACGCGAGAGGGCCGACGACGAGAGTGTCTTCAACT 5463  
 Db 857 GATTATATCAGCATTAACACGACGCGAGAGGGCCGACGACGAGAGTGTCTTCAACT 798  
 QY 5464 TTATCGGCTCCATCAGTCTATTAATTGTTCCGGGAAGTAGAGTAAGTAGTTCGGCA 5523  
 Db 797 TTATCGGCTCCATCAGTCTATTAATTGTTCCGGGAAGTAGAGTAAGTAGTTCGGCA 738  
 QY 5524 GTTAATAGTTGGCGCAAGCTGTGTCATTTGCTACAGGACATCGTGTCTACGCTGCTG 5583  
 Db 737 GTTAATAGTTGGCGCAAGCTGTGTCATTTGCTACAGGACATCGTGTCTACGCTGCTG 678

QY 5584 TTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAGCATCAAGCGAGTATCATATCCCCC 5643  
 Db 677 TTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAGCATCAAGCGAGTATCATATCCCCC 618  
 QY 5644 ATGTGTGCAAAAAAGCGGTAGTCTCTCGGTTCCGATGTTGTCAAGTAAGTTG 5703  
 Db 617 ATGTGTGCAAAAAAGCGGTAGTCTCTCGGTTCCGATGTTGTCAAGTAAGTTG 558  
 QY 5704 GCCGCAAGTTATTCATCATGTTATATGCGACATGACATATTTCTTACTGTACGCCA 5763  
 Db 557 GCCGCAAGTTATTCATCATGTTATATGCGACATGACATATTTCTTACTGTACGCCA 498  
 QY 5764 TCCGTAAGATGTTTCTGTGATCTGTGATGATCTCAACCAATCTTGTGAATATGCT 5823  
 Db 497 TCCGTAAGATGTTTCTGTGATCTGTGATGATCTCAACCAATCTTGTGAATATGCT 438  
 QY 5824 ATGCGGCGACGAGTTGCTCTTCCGCGGATATACGAGATATACCGCGCATATGAC 5883  
 Db 437 ATGCGGCGACGAGTTGCTCTTCCGCGGATATATACGAGATATATACCGCGCATATGAC 378  
 QY 5884 AGAAGCTTTAAAGTCTCATCATTTGAAAAAGCTTCTGCGGCGAAAACTCTCAAGATC 5943  
 Db 377 AGAAGCTTTAAAGTCTCATCATTTGAAAAAGCTTCTGCGGCGAAAACTCTCAAGATC 318  
 QY 5944 TTACCGGCTGTAGATTCAGTTCATGATGATACCATCTGTGACCAACCTGATCTTCAGCA 6003  
 Db 317 TTACCGGCTGTAGATTCAGTTCATGATGATACCATCTGTGACCAACCTGATCTTCAGCA 258  
 QY 6004 TCTTTTACTTTCACGAGGTTCTGTGGTGAGCAAAACGAAAGCAAAATCCGCAAA 6063  
 Db 257 TCTTTTACTTTCACGAGGTTCTGTGGTGAGCAAAACGAAAGCAAAATCCGCAAA 198  
 QY 6064 AAGGCAATTAAGGCGCACACGGAATGTGAATATCATCTTCTCTTTTCAATATAT 6123  
 Db 197 AAGGCAATTAAGGCGCACACGGAATGTGAATATCATCTTCTCTTTTCAATATAT 138  
 QY 6124 TGAACATTTATCAGGCTTATTTCTCATGACGCAATCATATTTTGAATGTATTAAGAA 6183  
 Db 137 TGAACATTTATCAGGCTTATTTCTCATGACGCAATCATATTTTGAATGTATTAAGAA 78  
 QY 6184 AATAACAAATATGAGGTTCCGCGCACATTTCCCGAAAGTCCAC 6229  
 Db 77 AATAACAAATATGAGGTTCCGCGCACATTTCCCGAAAGTCCAC 32

RESULT 6  
 AL959357 956 bp mRNA linear EST 27-NOV-2002  
 LOCUS AL959357 XGC-gastrula Silurana tropicalis cDNA clone Tgas127k19 5',  
 DEFINITION mRNA sequence.  
 ACCESSION AL959357  
 VERSION AL959357  
 KEYWORDS EST.  
 SOURCE AL959357.1 GI:25782952  
 ORGANISM Silurana tropicalis (western clawed frog)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Silurana.  
 REFERENCE 1 (bases 1 to 956)  
 AUTHORS Taylor R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.  
 Sanger Xenopus tropicalis EST project 2002  
 TITLE Sanger Xenopus tropicalis EST project 2002  
 JOURNAL Unpublished  
 COMMENT Contact: Taylor R  
 Sanger Centre  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: tropesanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: Tgas127k19.plksp6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.  
 location/Qualifiers

```

source
1. .956
/organism="Silurena tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="Tgas127k19"
/dev_stage="gastrolula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli xtl-blue"
/clone_lib="XGC-gastrolula"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5' end of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."

BASE COUNT      242 a      242 c      235 g      236 t      1 others
ORIGIN
Query Match      14.2% Score 882.8; DB: 9; Length 956;
Best Local Similarity 99.7%; Pred. No. 1.2e-164;
Matches 884; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

4464 GGGCGGCTTGGCGGTTTTCATAGGCTCCGCCCTGAGCAGCATCACAAAATCG 4523
|||||
65 GGGCGGCTTGGCGGTTTTCATAGGCTCCGCCCTGAGCAGCATCACAAAATCG 124
|||||
4524 AGCTCAAGTCAGAGGTGGCGAAACCGACAGACTATAAGATPACAGCGCTTCCCGC 4583
|||||
125 AGGCTCAAGTCAGAGGTGGCGAAACCGACAGACTATAAGATPACAGCGCTTCCCGC 184
|||||
4584 TGGAAAGTCCCTCGGGGCTCTCGTTCGACCGCTGCGGTTACCGGATACGTGCGC 4643
|||||
185 TGGAAAGTCCCTCGGGGCTCTCGTTCGACCGCTGCGGTTACCGGATACGTGCGC 244
|||||
4644 CTTCCTCCCTCGGGGAAAGCGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAGTTC 4703
|||||
245 CTTCCTCCCTCGGGGAAAGCGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAGTTC 304
|||||
4704 GGTTAGTGTCTTGGCTCCAAAGCTGGGCTGTGTGACAGACCCCGCTTACGCCGACCG 4763
|||||
305 GGTTAGTGTCTTGGCTCCAAAGCTGGGCTGTGTGACAGACCCCGCTTACGCCGACCG 364
|||||
4764 CTGGGCGCTTATCCGGTAACTATCGTCTTGAATCCCAACCGGTAAGACACGACTATGCGC 4823
|||||
365 CTGGGCGCTTATCCGGTAACTATCGTCTTGAATCCCAACCGGTAAGACACGACTATGCGC 424
|||||
4824 ACTGGACAGCAGCAGCTGTAAAGATTAAGAGACGAGATTAAGAGCGGTGCTACAGA 4883
|||||
425 ACTGGACAGCAGCAGCTGTAAAGATTAAGAGACGAGATTAAGAGCGGTGCTACAGA 484
|||||
4884 GTTCTTGAAGTGTGGCTTAAGTACGAGCTACACTAGAAAGACAGTATTGGTATCTGCGC 4943
|||||
485 GTTCTTGAAGTGTGGCTTAAGTACGAGCTACACTAGAAAGACAGTATTGGTATCTGCGC 544
|||||
4944 TCTGCTGAAGCAGTTCCTTCGGAAGAAAGTGTGATGCTTGAATCCCGCAAAACAAAC 5003
|||||
545 TCTGCTGAAGCAGTTCCTTCGGAAGAAAGTGTGATGCTTGAATCCCGCAAAACAAAC 604
|||||
5004 CACCGGTGTAGCGGTGTTTGTGTTTGAACAGAGATTAAGCGCAAAAAAAGG 5063
|||||
605 CACCGGTGTAGCGGTGTTTGTGTTTGAACAGAGATTAAGCGCGCAAAAAAAGG 664
|||||
5064 ATCTCAGAAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAAGAAACATC 5123
|||||
665 ATCTCAGAAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAAGAAACATC 724
|||||
5124 ACGTTAAGGATTTTGGTATGATATCAAAAAAGATCTTCACTAGATCTTTTAA 5183
|||||
725 ACGTTAAGGATTTTGGTATGATATCAAAAAAGATCTTCACTAGATCTTTTAA 784
|||||
5184 TTAATAAAGATTTTAATCAATCTAAAGATATATAGATAAACTGGTGCAGCTTA 5243
|||||
785 TTAATAAAGATTTTAATCAATCTAAAGATATATAGATAAACTGGTGCAGCTTA 844
|||||
5244 CCAATGCTTAATCAGTAGGACCACTATCTCAGCATCTGCTATTGCTTCAATCAGT 5303
|||||

```

```

|||||
Db      845 CCAATGCTTAATCAGTAGGACCACTATCTCAGCATCTGCTATTGCTTCAATCAGTANG 904
|||||
QY      5304 TGCCGAGTCCCGCTCGGTGATGATACATGATCGGAGGCGCTTAC 5350
|||||
Db      905 TGCTGACTCCCGCTGTTGATGATACATGATCGGAGGCGCTTAC 951
|||||

RESULT 7
Bg838279/c
LOCUS      Bg838279
DEFINITION Glycine claudestina cDNA clone Gc01_10e07, mRNA sequence.
ACCESSION      BG838279
VERSION      BG838279.1 GI:14204601
KEYWORDS      EST.
SOURCE      Glycine claudestina
ORGANISM      Glycine claudestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE      1 (bases 1 to 935)
AUTHORS      Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris
,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker
,N.A.
TITLE      Expressed Sequence Tags from Cold-Stressed Glycine claudestina
Seedlings
JOURNAL      Unpublished
COMMENT      Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
FEATURES
source
1. .935
/organism="Glycine claudestina"
/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_10e07"
/tissue_type="Leaves, stem"
/clone_lib="Gc01_AARC_ECORC_cold_stressed_glycine_claudestina"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
Site_2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed by
digest with Xho I/Eco RI and ligated to pBluescript."

BASE COUNT      243 a      213 c      227 g      238 t      14 others
ORIGIN
Query Match      14.0% Score 870.2; DB: 12; Length 935;
Best Local Similarity 98.1%; Pred. No. 3.6e-162;
Matches 874; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

5340 GAGAGGCTTACCATCTGGCCCGCACTGCTGCAATGATACCGGAGACCCAGCTTACCGGC 5399
|||||
928 GGGGGGCTTACCATCTGGCCCGCAKGTGCAATGATACCGGAGMCCAGCGCTMCGCGC 869
|||||
5400 TCCAGATTATCAGCAATTAACACGACGAGCCGGAAGGCGGAGCGGAGAGTGTCTGC 5459
|||||
868 YCCAGATTAATACAGCAATTAACACGAGCCGGAAGGCGGAGCGGAGAGTGTCTGC 809
|||||
5460 AACCTTATCCGCGCCATCCAGTCTATTAATTTGCGGGAGACTGAGTAAGTACTTC 5519
|||||
808 AWCCTTATCCGCGCCATCCAGTCTATTAATTTGCGGGAGAACTGAGTAAGTACTTC 749
|||||
5520 GCCAGTTAATAGTTTGGCAACGTTGTCATGTTGCAATGTTACAGAGCATCGTGTGTCACGCTC 5579
|||||

```

```

|||||
Db 748 GCCAGTATATAGTTGGCCACGTTGTCACATTCCTACAGGACATCGTGGTCCGCTC 689
QY 5580 GTCTGTTGGTATGGCTTATTCAGCTCCGGTTCCCAACGATCAGCGAGTTACATGATC 5639
Db 688 GTCTGTTGGTATGGCTTATTCAGCTCCGGTTCCCAACGATCAGCGAGTTACATGATC 629
QY 5640 CCCCATGTTGCAAAAAAGGGGTAGCTCCCTCGCTCCGATCGTTGTCAGAGTAA 5699
Db 628 CCCCATGTTGCAAAAAAGGGGTAGCTCCCTCGCTCCGATCGTTGTCAGAGTAA 569
QY 5700 GTT-GGCCGAGTGTATCAGTCATGTTATGAGCAGCTGCATTAATCTCTACTGTCA 5758
Db 568 GTTGGCCGCGAGTGTATCAGTCATGTTATGAGCAGCTGCATTAATCTCTACTGTCA 509
QY 5759 TGCATCCGTAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGAT 5818
Db 508 TGCCATCCGTAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGAT 449
QY 5819 AGTGTATCGGCGACCGAGTGTGCTGTGCTCCGCTCAATACGGATTAATCCGCGCAC 5878
Db 448 AGTGTATCGGCGACCGAGTGTGCTGTGCTCCGCTCAATACGGATTAATCCGCGCAC 389
QY 5879 ATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGCTTTCTGGGGGCAAAACTCTCAA 5938
Db 388 ATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGCTTTCTGGGGGCAAAACTCTCAA 329
QY 5939 GGATCTTACCGCTGTGATGATCCAGTGTGATGATACCCAGCTGTGACCCCAACTGATCT 5998
Db 328 GGATCTTACCGCTGTGATGATCCAGTGTGATGATACCCAGCTGTGACCCCAACTGATCT 269
QY 5999 CAGACTCTTTACTTTCACACGCGTTTGGGTGGGCAAAAAAGGAGGCAAAATGGCG 6058
Db 268 CAGACTCTTTACTTTCACACGCGTTTGGGTGGGCAAAAAAGGAGGCAAAATGGCG 209
QY 6059 CAAAAAGGAAATAGGCGACACGAGAAATGTGATACTCATCTCTCTTTTTCAT 6118
Db 208 CAAAAAGGAAATAGGCGACACGAGAAATGTGATACTCATCTCTCTTTTTCAT 149
QY 6119 ATTATGAGCAATTTATCAGGTTATGTCTCATGAGCGGATACATTAATTTGAATGATTT 6178
Db 148 ATTATGAGCAATTTATCAGGTTATGTCTCATGAGCGGATACATTAATTTGAATGATTT 89
QY 6179 AGAAAAATTAACAATAGGGTCCGCGACATTTCCCGGAAAGTGGCAC 6229
Db 88 AGAAAAATTAACAATAGGGTCCGCGACATTTCCCGGAAAGTGGCAC 38

RESULT 8
CD459092/c 918 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08e02_R Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08e02, mRNA sequence.
ACCESSION CD459092
VERSION CD459092.1 GI:31373832
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 918)
AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori,
J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL Unpublished
COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655

```

```

FEATURES
    source
        location/Qualifiers
            1..918
                /organism="Gibberella zeae"
                /mol_type="mRNA"
                /strain="DAOM 180378"
                /db_xref="taxon:518"
                /clone="Fg08_08e02"
                /lissue_type="Mycelium"
                /dev_stage="Asexual"
                /lab_host="E. coli DH108"
                /clone_lib="Fg08_AAFC_ECORC_Fusarium_graminearum_complex-s
                ubstrate"
                /note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
                XhoI; Fusarium graminearum grown on a complex plant
                substrate-- wheat leaves treated to remove most of the low
                molecular weight, water-soluble components."
BASE COUNT 247 a 208 c 321 g 239 t 3 others
ORIGIN
Query Match 13.8%; Score 862.6; DB 14; Length 918;
Best Local Similarity 99.4%; Pred. No. 1.2e-160;
Matches 884; Conservative 2; Mismatches 1; Indels 2; Gaps 2;
QY 5343 GGGCTTACATCTGGCCCCCAGTGCATATGA-TACCGCGAGACCCAGCTCACGGCTC 5401
Db 918 GGGCTTACATCTGGCCCCCAGTGCATATGA-TACCGCGAGACCCAGCTCACGGCTC 859
QY 5402 CAGATTATTCAGCAATTAACACGAGCGGAGGAGGCGAGAGTGGTCTCGCA 5461
Db 858 CAGATTATTCAGCAATTAACACGAGCGGAGGAGGCGAGAGTGGTCTCGCA 799
QY 5462 CTTTATCCGCTCATCCAGTCTATTAATTTGGCGGGAAGCTAGAGTAAGTTCGC 5521
Db 798 CTTTATCCGCTCATCCAGTCTATTAATTTGGCGGGAAGCTAGAGTAAGTTCGC 739
QY 5522 CAGTTAATAGTTGGCGCAACGTTTGGCATTTGCTACAGCATCTCGTGTGACCTCGT 5581
Db 738 CAGTTAATAGTTGGCGCAACGTTTGGCATTTGCTACAGCATCTCGTGTGACCTCGT 679
QY 5582 CGTTTGATAGGCTTCATTCAGCTCCGGTCCCAAGATCAAGGAGTTACATGATCC 5641
Db 678 CGTTTGATAGGCTTCATTCAGCTCCGGTCCCAAGATCAAGGAGTTACATGATCC 619
QY 5642 CCATCTTGTGCAAAAAAGCGGTTAGCTCTCTCGTCCGATCGTGTGACAAAGTAAAGT 5701
Db 618 CCATCTTGTGCAAAAAAGCGGTTAGCTCTCTCGTCCGATCGTGTGACAAAGTAAAGT 559
QY 5702 TGGCCGAGTGTATACATCA-TGGTTATGAGCAGACGACATTAATTCCTTACGTCCAG 5760
Db 558 TGGCCGAGTGTATACATCA-TGGTTATGAGCAGACGACATTAATTCCTTACGTCCAG 499
QY 5761 CCATCCGTAAAGATGTTTCTGTGACTGAGTACTCAACCAAGTCAATTCGAGAAATAG 5820
Db 498 CCATCCGTAAAGATGTTTCTGTGACTGAGTACTCAACCAAGTCAATTCGAGAAATAG 439
QY 5821 TGTATCGGCGACGAGTTGCTCTTGGCCGGCGTCAATACGGGATTAATACCGCGCACAT 5880
Db 438 TGTATCGGCGACGAGTTGCTCTTGGCCGGCGTCAATACGGGATTAATACCGCGCACAT 379
QY 5881 AGCAGAACTTTAAAGTCTCATCTTTGGAAGAGCTTTGGGGCGAAACTCTCAAG 5940
Db 378 AGCAGAACTTTAAAGTCTCATCTTTGGAAGAGCTTTGGGGCGAAACTCTCAAG 319
QY 5941 ATCTTACCGCTGTGAGATCCAGTGTGATTAACCACTGTGACCCCACTGATCTTCA 6000
Db 318 ATCTTACCGCTGTGAGATCCAGTGTGATTAACCACTGTGACCCCACTGATCTTCA 259
QY 6001 GCATCTTTTACTTACACAGGTTTCTGGGTAGCAAAAGAGAGGCAAAATATCCGCA 6060
Db 258 GCATCTTTTACTTACACAGGTTTCTGGGTAGCAAAAGAGAGGCAAAATATCCGCA 199

```

QY 6061 AAAAAGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTCTTTCAATAT 6120  
 Db 198 AAAAAGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTTTCAATAT 139  
 QY 6121 TATGAGCAATTTATGAGGGTATGTCTCATGAGGATCATATTTGAATGATTTAG 6180  
 Db 138 TATGAGCAATTTATGAGGGTATGTCTCATGAGGATCATATTTGAATGATTTAG 79  
 QY 6181 AAAAATAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCAC 6229  
 Db 78 AAAAATAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCAC 30  
 RESULT 9  
 CDS08PE5/c 986 bp mRNA linear HTC 07-JAN-2003  
 LOCUS Single read from an extremity of a full-length cDNA clone made from  
 DEFINITION Anopheles gambiae total adult females. 3-PRIME end of clone  
 FK0AA34AH01 of strain 6-9 of Anopheles gambiae (African malaria  
 mosquito).  
 ACCESSION BX022329  
 VERSION BX022329.1 GI:2571549  
 KEYWORDS HTC.  
 SOURCE Anopheles gambiae (African malaria mosquito)  
 ORGANISM Anopheles gambiae  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;  
 Anopheles.  
 REFERENCE 1 (bases 1 to 986)  
 Genoscope.  
 AUTHORS Direct Submission  
 TITLE Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)  
 - Web : www.genoscope.cns.fr  
 FEATURES  
 source  
 1. 986  
 /organism="Anopheles gambiae"  
 /mol\_type="mRNA"  
 /strain="6-9"  
 /db\_xref="taxon:7165"  
 /clone="FK0AA34AH01"  
 /plasmid="pME18S-FL"  
 /note="end : 3-PRIME"  
 BASE COUNT 264 a 233 c 237 g 252 t  
 ORIGIN  
 Query Match 13.7%: Score 851.4; DB 11; Length 986;  
 Best Local Similarity 97.4%: Pred.No.2e-158;  
 Matches 889; Conservative 0; Mismatches 16; Indels 8; Gaps 2;  
 QY 5325 GATTAAGTACGATACGAGGAGGCTTACCATCTGAGCCAGTCTGCAATGATACCGGAGA 5384  
 Db 986 GATTAAGTACGATACGAGGAGGCTTACCATCTGAGCCAGTCTGCAATGATACCGGAGA 927  
 QY 5385 CCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACACGACGCGGAGGCGGAGCG 5444  
 Db 926 CCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACACGACGCGGAGGCGGAGCG 867  
 QY 5445 CAGAAAGTGTCCTCAGCTTATTCGCGCTCATCCAGCTCATTAATTTGTTCCGGGAAGC 5504  
 Db 866 CAGAAAGTGTCCTCAGCTTATTCGCGCTCATCCAGCTCATTAATTTGTTCCGGGAAGC 807  
 QY 5505 TAGAGTAAGTAGTTCGCGCAGTAAATTAATTTGCGCAAGCTTGTCCATTGCTACAGCAT 5564  
 Db 806 TAGAGTAAGTAGTTCGCGCAGTAAATTAATTTGCGCAAGCTTGTCCATTGCTACAGCAT 747  
 QY 5565 CGTGGTTCACGCTCGTCTGTTGGTATGCTTCATTACGCTCCGTTCCCAACGATCAG 5624  
 Db 746 CGTGGTTCACGCTCGTCTGTTGGTATGCTTCATTACGCTCCGTTCCCAACGATCAG 687  
 QY 5625 GCGAGTTACATGATCCCGCAATGTTGTGCAAAAAAGCGGTTAGCTCTGGTCTCTCGAT 5684  
 Db 686 GCGAGTTACATGATCCCGCAATGTTGTGCAAAAAAGCGGTTAGCTCTGGTCTCTCGAT 627

QY 5685 CGTTGTGAGAAGTAGTTGGCGCAGTGTATATCATTGATGTTATGCGACACTGATAA 5744  
 Db 626 CGTTGTGAGAAGTAGTTGGCGCAGTGTATATCATTGATGTTATGCGACACTGATAA 567  
 QY 5745 TTCTCTTACTGTATGCTCATTCGCTATGATGATGTTTCTGTGACTGTGATGATCAACCA 5804  
 Db 566 TTCTCTTACTGTATGCTCATTCGCTATGATGATGTTTCTGTGACTGTGATGATCAACCA 507  
 QY 5805 GTCAATCTGAGAAGTAGTTGGCGCAGTGTATATCATTGATGTTATGCGACACTGATAA 5864  
 Db 506 GTCAATCTGAGAAGTAGTTGGCGCAGTGTATATCATTGATGTTATGCGACACTGATAA 447  
 QY 5865 TAATACCGCGCCACATAGCAGAACTTTAAAGTGCATCATCTTTGAAACGTTCTCGG 5924  
 Db 446 TAATACCGCGCCACATAGCAGAACTTTAAAGTGCATCATCTTTGAAACGTTCTCGG 387  
 QY 5925 GCGAAACTCTCAAGGATCTTACCGCTGTGATGATGATGATGATGATGATGATGATGAT 5984  
 Db 386 GCGAAACTCTCAAGGATCTTACCGCTGTGATGATGATGATGATGATGATGATGATGAT 327  
 QY 5985 ACCCACTGATCTTACGATCTTTTACTTTCACCAAGGTTCTGCGTGAAGCAAAACAGG 6044  
 Db 326 ACCCACTGATCTTACGATCTTTTACTTTCACCAAGGTTCTGCGTGAAGCAAAACAGG 267  
 QY 6045 AAGCAAAATGCGCCCAAAAGGGAATAGGCGGACGACGCAATGTTGAA-----TAC 6097  
 Db 266 AAGCAAAATGCGCCCAAAAGGGAATAGGCGGACGACGCAATGTTGAA-----TAC 207  
 QY 6098 TCATAGCTCTTC-CTTTTCAATATTTATGAAAGCAATTTATGATGATGATGATGATGAT 6156  
 Db 206 GATGATCTCTTCTGCTTTTCAATATTTATGAAAGCAATTTATGATGATGATGATGATGAT 147  
 QY 6157 GGATACATATTTGAATGATTTTGAATAAATAAATAGGAGTTCCGCGCATTTCC 6216  
 Db 146 GGATACATATTTGAATGATTTTGAATAAATAAATAGGAGTTCCGCGCATTTCC 87  
 QY 6217 CGAAAGTGCAC 6229  
 Db 86 CGAAAGTGCAC 74  
 RESULT 10  
 AL042026/c 841 bp mRNA linear EST 29-FEB-2000  
 LOCUS DKFP434E11.1 r1 434 (synonym: hless3) Homo sapiens cDNA clone  
 DEFINITION DKFP434E11.1 5', mRNA sequence.  
 ACCESSION AL042026  
 VERSION AL042026.1 GI:5421372  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 841)  
 AUTHORS Poustka A., Klein M., Mewes H.W., Gassenhuber J. and Wiemann S.  
 TITLE EST (Poustka, et al.)  
 JOURNAL Unpublished  
 COMMENT  
 Contact: Poustka A.J.  
 Department Lehrach  
 Max-Planck-Institute for Molecular Genetics  
 Inestrasse 73, 14195 Berlin, Germany  
 Tel: +49-30-84131623  
 Fax: +49-30-84131128  
 Email: poustka@mpg-berlin-dahlem.mpg.de  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by DKFZ (German Cancer Research Center,  
 Heidelberg/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 No SI sequence available.  
 This clone (DKFP434E11.1) is available at the RZPD in Berlin.



Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source

Location/Qualifiers

1. 841

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp434E111"

/tissue\_type="testis"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_id="434 (synonym: hlesj)"

/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

## BASE COUNT

226 a 192 c 202 g 221 t

ORIGIN

Query Match 13.5%; Score 841; DB 9; Length 841;

Best Local Similarity 100.0%; Pred. No. 2.3e-156;

Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

5355 TGGCCCGCAGTGTGCAATGATATACGCGGACGACCGGTCACCGGCTCCAGATTATACAGC 5414  
|||||  
841 TGGCCCGCAGTGTGCAATGATATACGCGGACGACCGGTCACCGGCTCCAGATTATACAGC 782  
5415 AATAAACCGACGACGCGGAGGCGCGAGCGAGAGTGTCTGCAACTTTATCCGCCCTC 5474  
|||||  
781 AATAAACCGACGACGCGGAGGCGCGAGCGAGAGTGTCTGCAACTTTATCCGCCCTC 722  
5475 CATCCAGTCTATTAATTGTGCGGGAGAGCTAGAGTAAGTATGTTCCAGTTAATAGTTT 5534  
|||||  
721 CATCCAGTCTATTAATTGTGCGGGAGAGCTAGAGTAAGTATGTTCCAGTTAATAGTTT 662  
5535 GCGCAAGCTGTGTCATGCTGCTACAGCATGCTGTGTCACAGCTGTGTTGGATAGGC 5594  
|||||  
661 GCGCAAGCTGTGTCATGCTGCTACAGCATGCTGTGTCACAGCTGTGTTGGATAGGC 602  
5595 TTCAATTCAGTCCGCTGCCAAGCATGAGGAGTATGATGATCCCATGTTGTGCA 5654  
|||||  
601 TTCAATTCAGTCCGCTGCCAAGCATGAGGAGTATGATGATCCCATGTTGTGCA 542  
5655 AAAAGCGGTAGCTCTCGTCTCGATCGTGTGTCAGAAAGTATGTTGCCCGAGTGT 5714  
|||||  
541 AAAAGCGGTAGCTCTCGTCTCGATCGTGTGTCAGAAAGTATGTTGCCCGAGTGT 482  
5715 ATCACTCATGTTTATGAGACATGCAATTTCTTACTGTGATGCGCATCCGTAAGT 5774  
|||||  
481 ATCACTCATGTTTATGAGACATGCAATTTCTTACTGTGATGCGCATCCGTAAGT 422  
5775 CTTTCTGTGACTGTGATGATCAACCAAGTCAATTCATGAGAAATGATGATGCGGAC 5834  
|||||  
421 CTTTCTGTGACTGTGATGATCAACCAAGTCAATTCATGAGAAATGATGATGCGGAC 362  
5835 GAGTTGCTCTGCGCGCGGCTCAATACGGGATTAATACGCGCATACAGAACTTTAA 5894  
|||||  
361 GAGTTGCTCTGCGCGCGGCTCAATACGGGATTAATACGCGCATACAGAACTTTAA 302  
5895 AGTGCTCATGTTGGAAGGCTCTGCGGCGGAGAAATCTCAAGACTTTACCGCTTT 5954  
|||||  
301 AGTGCTCATGTTGGAAGGCTCTGCGGCGGAGAAATCTCAAGACTTTACCGCTTT 242  
5955 GAGATCCAGTTCGATGTAACCACTCGTGACACCAAGTCAATTCATGAGATCTTTACCTT 6014  
|||||  
241 GAGATCCAGTTCGATGTAACCACTCGTGACACCAAGTCAATTCATGAGATCTTTACCTT 182  
6015 CACGAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAGGAATAG 6074  
|||||  
181 CACGAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAGGAATAG 122  
6075 GCGCAGCAGGAATGTTCAATACATCTCTCTCTTTTCAATTTATGAGCATTTA 6134  
|||||  
121 GCGCAGCAGGAATGTTCAATACATCTCTCTCTTTTCAATTTATGAGCATTTA 62  
6135 TCAGGCTATTGTCTCATGAGCGGATACATATTGATTTAGAAAAATTAACAAT 6194

|||||  
Db 61 TCAGGCTATTGTCTCATGAGCGGATACATATTGATTTGATTTAGAAAAATTAACAAT 2

QY 6195 A 6195

Db 1 A 1

## RESULT 11

CNS09072/c

LOCUS

DEFINITION

Single read from an extremity of a full-length cDNA clone made from

Anopheles gambiae total adult females. 3-PRIME end of clone

FK0A051AF10 of strain 6-9 of Anopheles gambiae (African malaria

mosquito).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Anopheles gambiae (African malaria mosquito)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;

Anopheles.

1 (bases 1 to 915)

Genoscope.

Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage ;

BP 191 91006 EVRI cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Location/Qualifiers

1. 915

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="6-9"

/db\_xref="taxon:7165"

/clone="FK0A051AF10"

/plasmid="PME18S-PL"

/note="end : 3-PRIME"

BASE COUNT

241 a 224 c 218 g 232 t

## ORIGIN

Query Match 13.4%; Score 834.4; DB 11; Length 915;

Best Local Similarity 99.3%; Pred. No. 4.7e-155;

Matches 838; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

5386 CCAGCTCAGCGGCTCCAGATTATACAGCAATTAACACGACCGGAGGCGGACGCG 5445  
|||||  
914 CCGCGCGCGGCGGCGCGAGATTATACAGCAATTAACACGACCGGAGGCGGACGCG 855  
5446 AGAAGTGTGCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCT 5505  
|||||  
854 AGAAGTGTGCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCT 795  
5506 AGAGTAAGTATGTCGCAAGTATAGTTTGGCAAGTGTGTTGCCATTGCTACAGCAATC 5565  
|||||  
794 AGAGTAAGTATGTCGCAAGTATAGTTTGGCAAGTGTGTTGCCATTGCTACAGCAATC 735  
5566 GTGGTTCACGCTGCTGTTGGTATGCTTCATTCAGCTCCGGTCCCAACGATCAAG 5625  
|||||  
734 GTGGTTCACGCTGCTGTTGGTATGCTTCATTCAGCTCCGGTCCCAACGATCAAG 675  
5626 CGAGTTACATATGCCCATGTTGTCGCAAAAAGGGTTAGTCTCTGGTCCCTCCGATC 5685  
|||||  
674 CGAGTTACATATGCCCATGTTGTCGCAAAAAGGGTTAGTCTCTGGTCCCTCCGATC 615  
5686 GTTGCAGAAAGTATGTTGGCGCAGTGTATACATCATGTTATGAGCACTGCATAT 5745  
|||||  
614 GTTGCAGAAAGTATGTTGGCGCAGTGTATACATCATGTTATGAGCACTGCATAT 555  
5746 TCTCTTACTGTCATCCATCCGTAAGATGCTTTCTGTGACGTGTGAGTACTCAACCAAG 5805  
|||||  
554 TCTCTTACTGTCATCCATCCGTAAGATGCTTTCTGTGACGTGTGAGTACTCAACCAAG 495

QY 5806 TCATTCGAGAAATAGTGTATGCGGCGACCGAGTTCCTTCCCGCGCTCAATACGGGAT 5865  
 |||||||  
 Db 494 TCATTCGAGAAATAGTGTATGCGGCGACCGAGTTCCTTCCCGCGCTCAATACGGGAT 435  
 |||||||  
 QY 5866 AATACCGCGCCACATACAGAACTTTAAAGTGTCTCATCTGAGAAAACGTTCTTGGGG 5925  
 |||||||  
 Db 434 AATACCGCGCCACATACAGAACTTTAAAGTGTCTCATCTGAGAAAACGTTCTTGGGG 375  
 |||||||  
 QY 5926 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACCTCGGCA 5985  
 |||||||  
 Db 374 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACCTCGGCA 315  
 |||||||  
 QY 5986 CCCAAGTATCTTACGATCTTTTACTTTCACAGCCCTTCTGGGTGAGCAAAAACAGGA 6045  
 |||||||  
 Db 314 CCCAAGTATCTTACGATCTTTTACTTTCACAGCCCTTCTGGGTGAGCAAAAACAGGA 255  
 |||||||  
 QY 6046 AGCGAAATGCGCGCAAAAAGGAAATAGGCGCAGCAGGAATGTGAATCTCATCTC 6105  
 |||||||  
 Db 254 AGCGAAATGCGCGCAAAAAGGAAATAGGCGCAGCAGGAATGTGAATCTCATCTC 195  
 |||||||  
 QY 6106 TTCTTTTCAATATTTATTAAGCATTTATCAGGTTATGTCTCATCAGCGATACATA 6165  
 |||||||  
 Db 194 TTCTTTTCAATATTTATTAAGCATTTATCAGGTTATGTCTCATCAGCGATACATA 135  
 |||||||  
 QY 6166 TTGAATGTATTTAGAAAATTAACAATAGGGGTTCCGCGCATTTCCCGAAAAGTG 6225  
 |||||||  
 Db 134 TTGAATGTATTTAGAAAATTAACAATAGGGGTTCCGCGCATTTCCCGAAAAGTG 75  
 |||||||  
 QY 6226 CCAC 6229  
 |||||  
 Db 74 CCAC 71

## RESULT 12

CD459085/c 872 bp mRNA linear EST 03-JUN-2003  
 LOCUS Fg08\_08d03\_R Fg08\_AAFc\_ECORC\_Fusarium\_graminearum\_complex-substrate  
 DEFINITION Gdbberella zeae cDNA clone Fg08\_08d03, mRNA sequence.  
 ACCESSION CD459085  
 VERSION CD459085.1 GI:31373825  
 KEYWORDS EST.

## SOURCE

ORGANISM Gdbberella zeae  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

## REFERENCE

AUTHORS Watson,R.J., Heys,R., Chapados,J., Couroux,P., Harris,L.J., Hattori  
 J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and  
 Tinker,N.A.

## TITLE

A cDNA library prepared from Fusarium graminearum grown on a  
 complex plant substrate

UNPUBLISHED  
 Contact: Watson, Robert.J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-Food Canada  
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,  
 CANADA

Tel: (613) 759-1655  
 Fax: (613) 759-1701  
 Email: watsonrj@agr.gc.ca.

Location/Qualifiers  
 1. 872

## FEATURES

Source  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg08\_08d03"  
 /issue\_type="Mycelium"  
 /dev\_stage="Asexual"  
 /lab\_host="E. coli DH10B"  
 /clone\_id="Fg08\_AAFc\_ECORC\_Fusarium\_graminearum\_complex-s  
 ubstrate"  
 /note="Vector: pBluescript II+; Site\_1: EcoRI; Site\_2:

Query Match 13.4%; Score 832.8; DB 14; Length 872;  
 Best Local Similarity 99.2%; Pred. No. 9,7e-155;  
 Matches 831; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
 BASE COUNT 235 a 194 c 210 g 228 t 5 others  
 ORIGIN

QY 5385 CCACAGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCCGGAAGGCGGAGCG 5444  
 |||||||  
 Db 872 CCACAGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCCGGAAGGCGGAGCG 813  
 |||||||  
 QY 5445 CAGAAGTGTCCTGCACTTATTCGGCTCCATTCATCTATTAATTTGTTCCGGGAAGC 5504  
 |||||||  
 Db 812 CAGAAGTGTCCTGCACTTATTCGGCTCCATTCATCTATTAATTTGTTCCGGGAAGC 753  
 |||||||  
 QY 5505 TAGAGTAAGTAGTTCGCGCAGTTAATAGTTTGGCGAACGTTGGTCCATGCTACAGCAT 5564  
 |||||||  
 Db 752 TAGAGTAAGTAGTTCGCGCAGTTAATAGTTTGGCGAACGTTGGTCCATGCTACAGCAT 693  
 |||||||  
 QY 5565 CGTGTGTACAGCTCGTCTGTTGGTATAGCTTATTCAGCTCCGTTCCCAAGCATCAAG 5624  
 |||||||  
 Db 692 CGTGTGTACAGCTCGTCTGTTGGTATAGCTTATTCAGCTCCGTTCCCAAGCATCAAG 633  
 |||||||  
 QY 5625 GCGAGTTACATGATNCCCATGTTGTGCAAAAAGCGGTTAAGTCCCTTCGGTCCCGAT 5684  
 |||||||  
 Db 632 GCGAGTTACATGATNCCCATGTTGTGCAAAAAGCGGTTAAGTCCCTTCGGTCCCGAT 573  
 |||||||  
 QY 5685 CGTTGTCAAGTAAGTGTGGCGCGAGTTCATCATCATGTTATGCGACACATGCTATA 5744  
 |||||||  
 Db 572 CGTTGTCAAGTAAGTGTGGCGCGAGTTCATCATCATGTTATGCGACACATGCTATA 513  
 |||||||  
 QY 5745 TTCTCTTACGTCTATGCCATCCGTAAGATCCTTTCTGTACATGCTGATCTCAACCA 5804  
 |||||||  
 Db 512 TTCTCTTACGTCTATGCCATCCGTAAGATCCTTTCTGTACATGCTGATCTCAACCA 453  
 |||||||  
 QY 5805 GTCATCTGAGAAATAGTGTATGGCGGACGAGTTCGTTGCCCGGCGTAATACGGGA 5864  
 |||||||  
 Db 452 GTCATCTGAGAAATAGTGTATGGCGGACGAGTTCGTTGCCCGGCGTAATACGGGA 393  
 |||||||  
 QY 5865 TAATACCGCGCCACATAGCAAACTTTAAAGTCTCATCATGTTGGAAGAGTTCCTGGG 5924  
 |||||||  
 Db 392 TAATACCGCGCCACATAGCAAACTTTAAAGTCTCATCATGTTGGAAGAGTTCCTGGG 333  
 |||||||  
 QY 5925 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTGTGC 5984  
 |||||||  
 Db 332 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTGTGC 273  
 |||||||  
 QY 5985 ACCCACTGATCTTACACATCTTTTACTTCCACAGGCTTTCGGGTGAGCAAAAACAGG 6044  
 |||||||  
 Db 272 ACCCACTGATCTTACACATCTTTTACTTCCACAGGCTTTCGGGTGAGCAAAAACAGG 213  
 |||||||  
 QY 6045 AAGGCAAAATGCGCGCAAAAAGGGAATAGGGGACACGGAATGTTGAATCTCATACT 6104  
 |||||||  
 Db 212 AAGGCAAAATGCGCGCAAAAAGGGAATAGGGGACACGGAATGTTGAATCTCATACT 153  
 |||||||  
 QY 6105 CTTCCTTTTCAATATTTATTAAGCATTTATCAGGTTATTTGCTCATGAGCGGATCAT 6164  
 |||||||  
 Db 152 CTTCCTTTTCAATATTTATTAAGCATTTATCAGGTTATTTGCTCATGAGCGGATCAT 93  
 |||||||  
 QY 6165 ATTGTAATGTATTTAGAAAATTAACAATAAGGGGTCCCGGCACTTCCCGGAAA 6222  
 |||||||  
 Db 92 ATTGTAATGTATTTAGAAAATTAACAATAAGGGGTCCCGGCACTTCCCGGAAA 35  
 |||||||

## RESULT 13

BM438950/c 854 bp mRNA linear EST 31-JAN-2002  
 LOCUS BM438950  
 DEFINITION IRLVr00491 Liver cDNA library Ictaiurus punctatus cDNA 5', mRNA  
 sequence.  
 ACCESSION BM438950

```

VERSION      BM438950.1  GI:18460672
KEYWORDS
SOURCE       Ictalurus punctatus (channel catfish)
ORGANISM     Ictalurus punctatus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
              Ictaluridae; Ictalurus.
              1 (bases 1 to 854)
REFERENCE    Peng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
AUTHORS      Transcription of channel catfish (Ictalurus punctatus): initial
TITLE        analysis of expressed sequence tags from the liver
JOURNAL      Unpublished
COMMENT      Contact: Liu ZJ
              The Fish Molecular Genetics and Biotechnology Laboratory,
              Department of Fisheries and Allied Aquacultures and Program of Cell
              and Molecular Biosciences
              Auburn University
              203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
              Tel: 334 844 4054
              Fax: 334 844 9208
              Email: zliu@acesag.auburn.edu
              Seq primer: M13 Reverse
              Location/Qualifiers
                source
                  1..854
                  /organism="Ictalurus punctatus"
                  /mol_type="mRNA"
                  /db_xref="taxon:7998"
                  /clone_lib="Liver cDNA library"
                  /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
                  Site_2: SalI"
BASE COUNT   226 a 198 c 221 g 209 t
ORIGIN
Query Match 13.2%; Score 824.6; DB 12; Length 854;
Best Local Similarity 99.3%; Pred. No. 4,1e-153;
Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
OY 5203 TCATCTAAGTATATATAGTAACTGGTGTGACAGTATACCAATGCTTATCAAGTGG 5262
DB 854 TCAATCTCAGCATATATAGTAACTGGTGTGACAGTATACCAATGCTTATCAAGTGG 795
OY 5263 GCACCTATCTCAGCATCTGTCTATTTCGTCATCCATAGTTCGTCGACCTCCCGCTG 5322
DB 794 GCACCTATCTCAGCATCTGTCTATTTCGTCATCCATAGTTCGTCGACCTCCCGCTG 735
OY 5323 TAGATACTACGATACGGAGGCGCTTACCATCTGGCCCGCAAGTGTGCAATGATACCGCA 5382
DB 734 TAGATACTACGATACGGAGGCGCTTACCATCTGGCCCGCAAGTGTGCAATGATACCGCA 675
OY 5383 GACCCAGCTCAGCGGCTCAGATTATCAGCAATAAACGACGCGGGAAGGCGCGAG 5442
DB 674 GACCCAGCTCAGCGGCTCAGATTATCAGCAATAAACGACGCGGGAAGGCGCGAG 615
OY 5443 CGCAGAAGTGCTCTGCAACTTATCCGCTCCATCAGTCAATTAATTTGTCGGGAA 5502
DB 614 CGCAGAAGTGCTCTGCAACTTATCCGCTCCATCAGTCAATTAATTTGTCGGGAA 555
OY 5503 GCTAGAGTAAGTAGTTCCGCACTTAATAGTTTGGCGAAGTGTGGCATTCCTACAGGC 5562
DB 554 GCTAGAGTAAGTAGTTCCGCACTTAATAGTTTGGCGAAGTGTGGCATTCCTACAGGC 495
OY 5563 ATCGGGGTGTCAGCTCGCTGTTGGTATAGGCTTCATTCAGTCCGCGTCCCAAGATGA 5622
DB 494 ATCGGGGTGTCAGCTCGCTGTTGGTATAGGCTTCATTCAGTCCGCGTCCCAAGATGA 435
OY 5623 AGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCCG 5682
DB 434 AGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCCG 376
OY 5683 ATCGTTGTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGACTGAT 5742
DB 375 ATCGTTGTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGACTGAT 316

```

```

OY 5743 AATTCCTTACTGTGATGCCATCCGATAGATGCTTTTCTGTAGCTGTGACTCAACC 5802
DB 315 AATTCCTTACTGTGATGCCATCCGATAGATGCTTTTCTGTAGCTGTGACTCAACC 256
OY 5803 AAGTCATTCGTGAAGATAGTATGCGCGGACGAGTGTGCTTCCCGGCTCAATACGG 5862
DB 255 AAGTCATTCGTGAAGATAGTATGCGCGGACGAGTGTGCTTCCCGGCTCAATACGG 196
OY 5863 GATTAATACGGCGCCACATAGCAGAACTTTAAAGTCTATCATATGGAAGAGTCTTGG 5922
DB 195 GATTAATACGGCGCCACATAGCAGAACTTTAAAGTCTATCATATGGAAGAGTCTTGG 136
OY 5923 GGGCGAAAACCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGAA-CCCACTCG 5981
DB 135 GGGCGAAAACCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGAA-CCCACTCG 76
OY 5982 TGCACCCAACTGATCTTCAGCATCTTTTACTTTACACACAGCTTTGCGTGAACAAAC 6041
DB 75 TGCACCCAACTGATCTTCAGCATCTTTTACTTTACACACAGCTTTGCGTGAACAAAC 16
OY 6042 AGGAAGCCAAATGCG 6056
DB 15 AGGAAGCCAAATGCG 1
RESULT 14
AL879792 993 bp mRNA linear EST 15-SEP-2002
LOCUS AL879792.XC-egg Silurana tropicalis cDNA clone Tegg019c18 3', mRNA
DEFINITION sequence.
ACCESSION AL879792
VERSION AL879792.1 GI:22900057
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
          Xenopodinae; Silurana.
          1 (bases 1 to 993)
REFERENCE Taylor, R., Ashurst, J.L., Cronling, M.D.R., Zorn, A.M. and Rogers, J.
AUTHORS Sanger Xenopus tropicalis EST project 2002
TITLE Unpublished
JOURNAL Contact: Taylor R
COMMENT Sanger Centre
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: trop@sanger.ac.uk
          Sanger Xenopus tropicalis EST project 2001
          TROPICALIS_SEQUENCE_ID: Tegg019c18.q1kT7
          Sequencing primer: 17
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Aaron M. Zorn.
          Location/Qualifiers
            source
              1..993
              /organism="Silurana tropicalis"
              /mol_type="mRNA"
              /db_xref="taxon:8364"
              /clone="Tegg019c18"
              /dev_host="Escherichia coli XL1-blue"
              /lab_host="XGC-egg"
              /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
              was oligo dt primed from 5ug of poly A+ RNA from egg.
              EcoRI-NotI cut cDNA was then ligated into pCS107 with
              EcoRI at the 5' end and NotI at the 3' end"
BASE COUNT 275 a 218 c 219 g 278 t
ORIGIN
Query Match 13.2%; Score 822.4; DB 9; Length 993;
Best Local Similarity 99.5%; Pred. No. 1.1e-152;
Matches 823; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 5403 AGATTATACGAATAAACACGCGGAGGCGCGGAGGCGGAGGAGTGTCTGCAAC 5462

```

Db	993	AGATTTATACAGCAATPAAACGACGACGCCGGAAGGGCCGACGGGCAAGAACTGGTCMTGCAAC	934
OY	5463	TTTATCCGCCCTCCATCCAGTCTATTAAATGTGTCCGGGAGCTAGATAGTACTGGCC	5522
Db	933	TTTATCCGCCCTCCATCCAGTCTATTAAATGTGTCCGGGAGCTAGATAGTACTGGCC	874
OY	5523	AGTTAATATGTTGGCAACGTTGTGTGGCCATTGCTACAGGCAATCGAGGTGTACCGTCGTC	5582
Db	873	AGTTAATATGTTGGCAACGTTGTGTGGCCATTGCTACAGGCAATCGAGGTGTACCGTCGTC	814
OY	5583	GTTTGGATGAGCTTCATTCACTCACTCCGCTCCGATTTCCCAACGATCAAGGGGAGTTACATGATCCC	5642
Db	813	GTCGTGGATGAGCTTCATTCACTCACTCCGCTCCGATTTCCCAACGATCAAGGGGAGTTACATGATCCC	754
OY	5643	CATGTGTGCAAAAAAGGGGTTTACTGCTCTCTGGGTCGCCATTCGTTGTCACAAGTAAGTT	5702
Db	753	CATGTGTGCAAAAAAGGGGTTTACTGCTCTCTGGGTCGCCATTCGTTGTCACAAGTAAGTT	694
OY	5703	GGCCGCACTGTTATACACTCATGATGTTATGAGCAGCACTGCATATTTCTTACTGTCATGCC	5762
Db	693	GGCCGCACTGTTATACACTCATGATGTTATGAGCAGCACTGCATATTTCTTACTGTCATGCC	634
OY	5763	ATCCGTAAGATGCTTTTCTGTGACTGGTGAATCTCAACCAAGTCATTTCTGAAATAGTG	5822
Db	633	ATCCGTAAGATGCTTTTCTGTGACTGGTGAATCTCAACCAAGTCATTTCTGAAATAGTG	574
OY	5823	TATCGCGGACGAGTGTGCTCTTCGCCGCGCTCAATACGGGATTAATACCGGCCACATAG	5882
Db	573	TATCGCGGACGAGTGTGCTCTTCGCCGCGCTCAATACGGGATTAATACCGGCCACATAG	514
OY	5883	CAGAACTTTAAATAGTGCATCATTTGGAAAAACGTTCTTCGGGGGCAAACTCTCAAGAT	5942
Db	513	CAGAACTTTAAATAGTGCATCATTTGGAAAAACGTTCTTCGGGGGCAAACTCTCAAGAT	454
OY	5943	CTTACCGGTGTGAGATCTCAGTTGATGATTAACCCACTCGTGACCACCACTGATCTTGACG	6002
Db	453	CTTACCGGTGTGAGATCTCAGTTGATGATTAACCCACTCGTGACCACCACTGATCTTGACG	394
OY	6003	ATCTTTTACTTTCACACAGCGTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAA	6062
Db	393	ATCTTTTACTTTCACACAGCGTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAA	334
OY	6063	AAAGGAAATAGGGCGACACGGAATGTTGAATCTCACTTCTCTTTTCAATATTA	6122
Db	333	AAAGGAAATAGGGCGACACGGAATGTTGAATCTCACTTCTCTTTTCAATATTA	274
OY	6123	TTGAAGCATTTATCAGGGTTATGTCTCATGAGCGGATCATATTTGAATGATTTAGAA	6182
Db	273	TTGAAGCATTTATCAGGGTTATGTCTCATGAGCGGATCATATTTGAATGATTTAGAA	214
OY	6183	AAATPAAACAAATAGGGGTTCCGGCCACACTTCCCGGAAATGTCAC	6229
Db	213	AAATPAAACAAATAGGGGTTCCGGCCACACTTCCCGGAAATGTCAC	167
RESULT 15			
CD458333/c			
LOCUS	CD458333	870 bp	mRNA linear EST 03-JUN-2003
DEFINITION	Fg08_09e07_R Fg08_AAF_CECORC_Fusarium_graminearum_complex_substrate		
ACCESSION	Gibberella zeae cDNA clone Fg08_09e07, mRNA sequence.		
VERSION	CD458333		
KEYWORDS	CD458333..1 GI:31373073		
SOURCE	EST.		
ORGANISM	Gibberella zeae		
	Gibberella zeae		
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
	Hypocreomycetidae; Hypocerales; Nectriaceae; Gibberella.		
REFERENCE	1 (bases 1 to 870)		
AUTHORS	Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori-I,		
	J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and		
	Tinker,N.A.		
TITLE	A cDNA library prepared from Fusarium graminearum grown on a		
	complex plant substrate		

Journal Comment	Unpublished Contact: Watson, Robert J. Eastern Cereal and Oilseed Research Centre Agriculture and Agri-Food Canada Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA Tel: (613) 759-1655 Fax: (613) 759-1701 Email: watsonrj@agr.gc.ca.
FEATURES	Location/Qualifiers
source	1..870
	/organism="Gibberella zeae"
	/mol_type="mRNA"
	/strain="DAOM 180378"
	/db_xref="taxon:5518"
	/clone="Fg08_09e07"
	/tissue_type="Mycelium"
	/dev_host="Asexual"
	/lab_host="E. coli DH10B"
	/clone_lib="Fg08_AARC_ECORC_Fusarium_graminearum_complex_s
	ubstrate"
	/note="vector: pbluescript II+, Site.1: EcoRI, Site.2: XhoI; Fusarium graminearum grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."
BASE COUNT	239 a 195 c 203 g 230 t 3 others
ORIGIN	
Query Match	13.1%: Score 816.8; DB.14; Length 870;
Best Local Similarity	99.4%: Pred. No.1.5e-151;
Matches 837; Conservative	3; Mismatches 0; Indels 2; Gaps 2;
OY	5389 CGCTACCGGGCTCCAGATTATTCAGCATATAACGACGACGGCGGAGGGCCGAGCGAGA 5448
Db	870 CGCTCMCCGGGCTCCAGATTATTCAGCATATAACGACGACGGCGGAGGGCCGAGCGAGA 811
OY	5449 AGTGCTCTCGAAGTTATATCCGCTCCATCCAGTCATTAATGCTTCCGGGAAGCTAGA 5508
Db	810 AGTGCTCTCGAAGTTATATCCGCTCCATCCAGTCATTAATGCTTCCGGGAAGCTAGA 751
OY	5509 GTAAGTAGTTGCGCAGTAACTAGTTTGCAGCAAGCGTTGTGCCATTCGTACAGGCATGTG 5568
Db	750 GTAAGTAGTTGCGCAGTAACTAGTTTGCAGCAAGCGTTGTGCCATTCGTACAGGCATGTG 691
OY	5569 GTGTACAGCTGCTGTTTGGTAGGCTTCATTCAGTCGCGGTTCCACAGATCAAGCGGA 5628
Db	690 GTGTACAGCTGCTGTTTGGTAGGCTTCATTCAGTCGCGGTTCCACAGATCAAGCGGA 631
OY	5629 GTTACAGATCCCCCATGTTGTGCAAAAAACGGTTAGTCCTCGGTCCTCCGATGTT 5688
Db	630 GTTACAGATCCCCCATGTTGTGCAAAAAACGGTTAGCT-CTTCGGTCTCTCGATGTT 572
OY	5689 GTTCAG-AAGTAGTTGGCCGCGAGTGTATCATCTCATGTTATGCGACACTGCATAATTC 5747
Db	571 GTTCAGAAAGTAGTTGGCCGCGAGTGTATCATCTCATGTTATGCGACACTGCATAATTC 512
OY	5748 TCTTACAGTATGCGATCCGTAAGATGCTTTCGTGACGTGAGATCAACCAAGTC 5807
Db	511 TCTTACAGTATGCGATCCGTAAGATGCTTTCGTGACGTGAGATCAACCAAGTC 452
OY	5808 ATTCTGAGAAATGATGCGGACCGCAGTGTCTTTCGCCGCGCTCAATACGGAGTAA 5867
Db	451 ATTCTGAGAAATGATGCGGACCGCAGTGTCTTTCGCCGCGCTCAATACGGAGTAA 392
OY	5868 TACCGCGGCACATGCGAAGACTTTAAAAAGTCTCATATTGAGAAAGCTTCCGGGGCG 5927
Db	391 TACCGCGGCACATGCGAAGACTTTAAAAAGTCTCATATTGAGAAAGCTTCCGGGGCG 332
OY	5928 AAAACTGTCAAGATCTTACCGGTGTGAGATCCAGTTCATGTAAACCACTGTGAC 5987
Db	331 AAAACTGTCAAGATCTTACCGGTGTGAGATCCAGTTCATGTAAACCACTGTGAC 272
OY	5988 CAACGTGATCTTACACATCTTTACTTTCACACAGGTTTCGGGGAGCAAAAACAGGAG 6047

```

Db      271  CAACTGATCTTCAGCATCTTTACTTCCACGCGTCTCTGSGTGAGCAAAAACAGGAAG 212
Qy      6048 GCAAAATGCCGCAAAAAGGGAATAGGGCGACGCAAAATGTGAATACTCTACTCTT 6107
Db      211  GCAAAATGCCGCAAAAAGGGAATAGGGCGACGCAAAATGTGAATACTCTACTCTT 152
Qy      6108 CCTTTCAATATATATGAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATAT 6167
Db      151  CCTTTTCATATATATGAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATAT 92
Qy      6168 TGAATGTATTTAGAAAAATTAACAATAAGGGTTCGCGCACATTTCCCGAAAAAGTGC 6227
Db      91  TGAATGTATTTAGAAAAATTAACAATAAGGGTTCGCGCACATTTCCCGAAAAAGTGC 32
Qy      6228 AC 6229
Db      31  AC 30

```

Search completed: September 17, 2003, 17:25:02  
 Job time : 11665 secs

**THIS PAGE BLANK (USPTO)**